

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 08:11:50 ; Search time 4658 Seconds
(without alignments)
3024.149 Million cell updates/sec

Title: US-09-787-737-2

Perfect score: 1751

Sequence: 1 MSSSNKNWPMFKSPKNNN.....TDFGVYQPLQNGAIYILI 325

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2.1/USPTO.spool.h/US09787737/runat_01092004_112454_23595/app_query.fasta_1.519
-Db=GenEmbl -OEMT-fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US09787737 @CGN 1.1 3731 @runat_01092004_112454_23595 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DAV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htg.hum.*
40: em.htg.mus.*
41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1751	100.0	1214	6	E52177	B52177 Homeobox ge
2	1751	100.0	1214	6	BD012130	BD012130 Homeobox
3	1586	90.6	84510	8	AB006698	AB006698 Arabidops
4	645.5	36.9	1518	6	E52178	B52178 Homeobox ge
5	645.5	36.9	1518	6	BD012131	BD012131 Homeobox
6	643.5	36.8	1500	8	AK118501	AK118501 Arabidops
7	576	32.9	114144	8	U78721	U78721 Arabidops
8	420.5	24.0	918	8	BT005296	BT005296 Arabidops
9	394.5	22.5	807	8	BT006232	BT006232 Arabidops
10	394.5	22.5	996	8	AK117910	AK117910 Arabidops
11	369.5	21.1	159878	2	AC124143	AC124143 Oryza sat
12	369.5	21.1	182868	2	AC121360	AC121360 Oryza sat
13	347.5	19.8	182681	8	AP003760	AP003760 Oryza sat
14	338	19.3	1282	8	AK063503	AK063503 Oryza sat
15	338	19.3	1378	8	AK073232	AK073232 Oryza sat
16	337.5	19.3	81701	8	AB006706	AB006706 Arabidops
17	328.5	18.8	1367	6	AX105291	AX105291 Sequence
18	312	17.8	1449	8	AK111000	AK111000 Oryza sat
19	312	17.8	1602	8	AK063262	AK063262 Oryza sat
20	307.5	17.6	89334	8	ATAC009327	ATAC009327 Arabidops
21	299.5	17.1	1247	8	AK107305	AK107305 Oryza sat
22	280.5	16.0	67448	8	AP005249	AP005249 Oryza sat
23	237.5	13.6	152365	8	AP005167	AP005167 Oryza sat
24	236	13.5	162711	8	AC137070	AC137070 Genomic s
25	213	12.2	103125	8	AC005560	AC005560 Arabidops
26	212.5	12.1	1326	8	AY085832	AY085832 Arabidops
27	209.5	12.0	1117	8	AY162209	AY162209 Antirrhin
28	200	11.4	1209	8	AY086385	AY086385 Arabidops
29	199.5	11.4	1227	8	ATH012310	ATH012310 Arabidops
30	194.5	11.1	1137	8	AF481951	AF481951 Petunia x
31	193	11.0	88149	8	AC006201	AC006201 Arabidops
32	192.5	11.0	131346	8	CNS08CDT	CNS08CDT Oryza sat
33	192.5	11.0	147501	2	CNS0954R	CNS0954R Oryza sat
34	192.5	11.0	161808	8	CNS08CDM	CNS08CDM Oryza sat
35	192	11.0	1180	8	AK102203	AK102203 Oryza sat
36	189.5	10.8	76444	8	AP006099	AP006099 Lotus cor
37	187.5	10.7	1338	6	AX105285	AX105285 Sequence
38	187	10.7	741	6	AX105301	AX105301 Sequence
39	187	10.7	1035	6	LES538329	LES538329 lycopersi
40	187	10.7	144305	2	AC137078	AC137078 Arabidops
41	186.5	10.7	74968	8	AB019230	AB019230 Arabidops
42	183	10.5	134982	8	AP003279	AP003279 Oryza sat
43	183	10.5	150976	8	AP003442	AP003442 Oryza sat
44	178.5	10.2	749	8	AY150812	AY150812 Arabidops
45	178.5	10.2	74494	8	AC007171	AC007171 Arabidops

ALIGNMENTS

RESULT 1

E52177
LOCUS E52177 1214 bp DNA linear PAT 31-JAN-2002
DEFINITION Homeobox gene encoding protein participating in differentiation.
ACCESSION E52177
VERSION E52177.1 GI:18629640
KEYWORDS JP 2001029081-A/1.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1214)
REFERENCE
AUTHORS Kakimoto, T.
TITLE Homeobox gene encoding protein participating in differentiation
JOURNAL Patent: JP 2001029081-A 1 06-FEB-2001;
SUNTORY LTD, NIPPON PAPER IND CO LTD
OS Arabidopsis thaliana (thale cress)
COMMENT
PN JP 2001029081-A/1
PD 06-FEB-2001
PF 22-JUL-1999 JP 1999207995
PR TATSUO KAKIMOTO
PI C12N15/09, A01H1/00, A01H5/00, C07K14/42, C12N1/15, C12N1/19, PC
PC C12N1/21, C12N5/10,
PC C12P21/02, C12N15/00, C12N5/00
CC
FH
FT
Location/Qualifiers
i. .1214
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-157 Length: 1214
Score: 1751.00 Matches: 325
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-787-737-2 (1-325) x E52177 (1-1214)
QY 1 MetSerSerAsnLysAsnTyrProSerMetPheLysSerLysProCysAsnAsn 20
DB 36 ATGTCTCTCCCAACAAATAATGGCCAGCATGTTCAATCCAAACCTTGCAACAATAAT 95
QY 21 HisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
DB 96 CATCATCATCAACATGAATCGATCTCATCTTACATGCACTACTCTAATGTCAACCTA 155
QY 41 SerSerSerPheSerAspArgLysProAspProLysProAspTyrAsnProLysPro 60
DB 156 TCATCTCTCTTCTCTCAGATCGGATACCATCTTCAATCCAGATGGAATCTTAACCG 215
QY 61 GluGlnLeuArgLysLeuGluSerPheAsnSerGlyThrLeuAsnProProArgGlu 80
DB 216 GAGCAGATTAGATACATCAATCAATCTTCAATCCGGTACTAATTAACCCACCTAGAG 275
QY 81 GluLeuGlnArgLysLeuGluLeuGluGluGluGluGluGluGluGluGluGluGlu 100
DB 276 GAGATTCAAGAATCCGATCCGGCTTCAAGATATGTCCTCAATCCGGTACCAACCGT 335
QY 101 PheTyrTyrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120
DB 336 TTTTACTGGTTTCAAAACCGGAATCTCGAGCAAAACCAACAGCTTCGTTCATCACAA 395
QY 121 SerProLysMetSerLysLysAspLysThrValLysProSerThrAspAlaAspHisCys 140
DB 396 AGCCCTAAATGTCAAGAAGACAAAGACGATTTATCTCTAGTACTGACCTGATCATTTGT 455

QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValVal 160
DB 456 TTTGGTTTGTATACAGAAACCGGATATATCCGGTTCAAAACAATGAGTTGGTGA 515
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180
DB 516 ACCGAACGGCGGTTTCTATTCCGGTTCATATGATCCGAGCGCTCTCAATCAGCG 575
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
DB 576 TTTGGTTTGGCGATTTTCTTACCGGTGGTAACGGAAGAGGATGCAATCTCTACC 635
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysLeuProAlaLeuAsn 220
DB 636 GTTATACGGCGTTAAATTTGGAGCTAACGAAATTTTATAAAATTCGGCGATCAAT 695
QY 221 LeuTyrGlyGlyAspGlyAsnGlyGlyLysAsnCysPheProProLeuThrValProLeu 240
DB 696 TTATACGGCGGAGATGGAAATGGCGTGGAAATTTGTTTCTCTCTTGACTGTTCATTA 755
QY 241 ThrLeuAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGly 260
DB 756 ACCATCAATCAATCTCAAGAAACAGAGATGATGATTAACGAGATGCTATCGAAGTAGTG 815
QY 261 AspAsnValTyrProValArgMetThrValPheLeuAsnGluMetProLysGluValVal 280
DB 816 GATTAATGTTTATCCGGTGGAGATGACGGTGTATTAAACGAGATGCTATCGAAGTAGTG 875
QY 281 SerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuLeuAsnSerPhe 300
DB 876 TCTGGATTATTCAACGTTAAGCGAGCTTTTCGGAACGATGCCGTTTGTATCAACTCGTTT 935
QY 301 GlyGlnProLysLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyVala 320
DB 936 GGCCAGCTATTCTTACGATGAATTTGGTGTACTTATCACTTATCACTTCCAAATGGCGCA 995
QY 321 IleTyrTyrLeuLeu 325
DB 996 ATCTATTATCTTATT 1010
RESULT 2
BD012130
LOCUS Homeobox gene coding for protein participating in differentiation.
DEFINITION Homeobox gene coding for protein participating in differentiation.
ACCESSION BD012130
VERSION BD012130.1 GI:22092319
KEYWORDS WO 0107618-A/1.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1214)
REFERENCE
AUTHORS Kakimoto, T.
TITLE Homeobox gene coding for protein participating in differentiation
JOURNAL Patent: WO 0107618-A 1 01-FEB-2001;
SUNTORY LTD, TATSUO KAKIMOTO
OS Arabidopsis thaliana (thale cress)
COMMENT
PN WO 0107618-A/1
PD 01-FEB-2001
PF 21-JUL-2000 WO 2000JP004904
PR 22-JUL-1999 JP 99P 207995
PI TATSUO KAKIMOTO
PC C12N15/29, C12N1/21, C07K14/415, C12P21/02, C12N5/10, A01H5/00// PC
PC C12N15/29, C12N1/91, (C12N1/21, C12R1.01), (C12P21/02, C12R1.01) CC
Nucleotide sequence coding for protein participating in differentiation
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
i. .1214
/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores: 1.11e-157 Length: 1214
Pred. No.: 1751.00 Matches: 325
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-787-737-2 (1-325) x BD012130 (1-1214)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20
DB 36 ATGTCCTCCTCAACAAAATGGCCAAAGCATGTTCAAATCCAAACCTTGCACAAATAT 95
QY 21 HisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
DB 96 CATCATCATCAACATGAATCGATACCTCCATCTTACATGCACACTCTAATTGCAACCTA 155
QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60
DB 156 TCATCTTCCTTTTCCCTCAGATCGGATACCGATCCTAAACCGAGATCGAATCCTAAACCG 215
QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80
DB 216 GAGCAGATTAGGATACGATCGAATCAATCTCAATTCGGTACTATTAAACCCACCTAGAG 275
QY 81 GluLeuGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
DB 276 GAGATTCAAGAATCCGGATCCGGCTTCAAGATATGTCATCGTACGTCGACGCAACGTCG 335
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
DB 336 TTTTACTGGTTTCAAAACCGGAATCTCGAGCAAAACACAAAGCTTCGTGTTTCATCAAAA 395
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAsnHisCys 140
DB 396 AGCCCTAAATGTCAAGAGGACAGACGGTTATTCTAGTACTACCGTCGATCATTTGT 455
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160
DB 456 TTTTGGTTTGTGTTAAACCAAGACCGGATTATATCCGGTTCAAAACCAATGATGTTGGTA 515
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180
DB 516 ACCGAACCGCGCGGTTTCTATTTCGGGTTCAATAATGATCCGAGCGCTGCTCAATCAGCG 575
QY 181 PheGlyPheGlyAspPheValValProValThrGluGluGlyMetAlaPheSerThr 200
DB 576 TTTTGGTTTGGCGATTGTTGTTACCGGTGGTTAAACGAGAGAGGATGCGATCTCTAC 635
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsn 220
DB 636 GTTAATAACCGCGTTAAATTTGGAGACTACGAAAAATTTTGATAAAATTCGCGCATCAT 695
QY 221 LeuTyrGlyGlyAspGlyAsnGlyGlyAsnCysPheProProLeuThrValProLeu 240
DB 696 TTATACGGCGGAGATGGAATGGCGTGGAAATGTTTTCCTCCCTTGAATGTTCCATTA 755
QY 241 ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGly 260
DB 756 ACCATCATCATCTCAAGAAAAACGAGATGAGGATTTATCCGGTGGTGAAGACGTCGGA 815
QY 261 AspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValVal 280
DB 816 GATAATGTTTATCCGGTGAGAAATGACGGTGTGTTTATTACGAGATCGCTATCGAATGATG 875
QY 281 SerGlyLeuPheAsnValLysAlaPheGlyAsnAspAlaValLeuIleAsnSerPhe 300
DB 876 TCTGGATTATTCAAGGTCAGCTTTTCGAAACGATCGCGTTTTCATCAACTCGTTT 935

QY 301 GlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAla 320
DB 936 GGCCAGCCTATTCTTACAGATGAATTTGGTTACTTATCAACCTCTCCAAATGGCGCA 995
QY 321 IleTyrTyrLeuIle 325
DB 996 ATCTATTATCTTATT 1010

RESULT 3
AB006698

LOCUS 84510 bp DNA linear PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MCL19.
ACCESSION AB006698 BA000015
VERSION AB006698.1 GI:2351063

KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

AUTHORS

Kotani,H., Nakamura,Y., Sato,S., Kaneko,T., Asamizu,E., Miyajima,N.
and Tabata S.

TITLE

Structural analysis of Arabidopsis thaliana chromosome 5. II.

Sequence features of the regions of 1,044,062 bp covered by

thirteen physically assigned P1 clones

DNA Res. 4 (4), 291-300 (1997)

98069011

9405937

2 (bases 1 to 84510)

Nakamura,Y.

Direct Submission

Submitted (22-AUG-1997) Yasukazu Nakamura, Kazusa DNA Research

Institute, Department of Plant Gene Research; 1532-3, Yana,

Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,

Tel:81-438-52-3935, Fax:81-438-52-3934)

Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see

http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MCL19

Genes with similarity to proteins in the databases are described in

'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,

http://combio.ornl.gov/Grail-1.3/),

GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of

Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and

SplicePredictor (Volker Brendel, Stanford University,

http://gremli.2001.iastate.edu/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE

(Sean Eddy, Washington University School of Medicine, St. Louis,

http://genome.wustl.edu/eddy/tRNAscan-SE/).

This sequence may not be the entire insert of this clone. It may be

shorter because we remove overlaps between neighboring submissions.

The 5' clone is K15122 and the 3' clone is MDE13.

Location/Qualifiers

1. 84510

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia"

/db_xref="taxon:3702"

/chromosome="5"

/clone="MCL19"

/clone_lib="Mitsui P1"

join(1437..1535,1857..1944,2030..2139,2260..2323,

2912..2976,3055..3120,3321..3434)

/note="gene id:MCL19.1"

sp|Q38903"

/codon_start=1

/evidence=not_experimental

/product="Rac-like gtp binding protein ARAC3"

FEATURES
source

CDS

/protein id="BAB08242.1"
/db_xref="GI:9757717"
/translation="MSTARFKICVTVGAVGKTCMLISYTSNTPTDYVPTVDFNFS
ANVVDGTVNLGMDTAGQBYNLRPLSYRGADVFLLAFLSLIKASYENHKKWLP
ELKHVAPGIPVILVTKLDRDKQKLDKHPGASITTAQGBELKMGIGAVRYLECS
KIQQNVKAVDFDAIRVALRPPAKKKIKKTLKRKICFEL"
/note="contains similarity to homeodomain transcription
factor
gene id:MCL19.2"
/codon_start=1
/evidence=not_experimental
/protein id="BAB08243.1"
/db_xref="GI:9757718"
/translation="MSSSNKNMPSMFKSKPCNNHHQHEIDTPSMYHNSLSSFF
SDRIPDPKRWPKPEQIRILIESFNSGTINPPREIRIRIQEYQIGDANVYF
WPNRKRKAKHLRVHHSKPKSKDKTVLPSTADHCFGVNQETGLYPVQNNELV
TFAGHLFPVNDPSAAQAFGDFVVPVTEEGNAFVNVGNLETFNENFDKIPA
INLYGDGNGGNCFFPLIVPTINOSQKLLISK"
/join(9000..10120,10206..10271,10355..10499,10600..10818,
10909..11071,11182..11377,11457..11568)
/note="gene id:MCL19.3"
/codon_start=1
/evidence=not_experimental
/product="CRN (crooked neck) protein"
/protein id="BAB08244.1"
/db_xref="GI:9757719"
/translations="MTGDAEVLKPLRTTRVKNKTPAPVQITARQILREARERQAEIR
PKQKITDSELSYDLRRKEPEDQIRRARWNIQVVKVAKVEEQMDYARASVME
RALEGYRHTLWVKAPEMKNKFNVNARVNDRSVTLPLPRVDLWEKVIYVEEKLIG
NVTGAKQIFERNWNSPDQKAWLCFKIPELRYNEIRASRIYERVLCHKPKSAFIRY
AKFEMRGQGVKLARVEYRAVDKLANDEAEILVFSFAFEERCKEVEKARTIYKPA
LDHIRKRAEELYKFAVEKQYQKEGIEDAIVGKRFYEYDEVEKPNLNYSDFY
VLYEYKNDKRIEYERAIANVPAQEKRFQWRYIYMINVLYEEIEKDVETR
DVYRECKLIPHTKFSFAKTLAAVEYIRQLMTGARQILGNAIGKAPKVKFKFYI
EMELKLVNDRCKLIERELWSPENCYAWNVAEPISEATERARAPELAI SOPA
LMPPELLWIKITIDFELSEGFEKTRALYERLORTCKEYKWIWISFAKFEASAKHEDG
IKSARVIFRANTYKIDTPELEBESRATILLEDMLNMTGEGELGSDVSVOGSKLPKLLK
KRWTSREDSYERYFDYLPFEESGTNNLKIIEAAKWKLLKGEVCV"
/complement(join(111995..12297,12529..12809,12975..13130,
13269..13404,13629..13736,13846..14052))
/note="gene id:MCL19.4"
/codon_start=1
/evidence=not_experimental
/product="myrosinase-binding protein-like; jasmonate
inducible protein-like"
/protein id="BAB08245.1"
/db_xref="GI:9757720"
/translation="MSKVIYAGSGNGKFDGQVYEGTVGVKLVGDEHGIYVL
KIQYKNGDVVVEHGRARGTHITETFEVKCPDEYITSIWTCGRNDHHRKSTGGRT
FEQLTADKYDINLAARMGTPMYNTVSELOFTKSHGRTSEQFRPGTGGSGVSWDD
GAYDRSLKLCVGEDAHCVSSEYFHYVKGNDRIITHCGKDSKEHDKGFISSLTKTSM
NRSEKFKGPVGFQLEAKGDFDKIVGFRGRSSVNRINALGANFAVAVVPPVKKNAK
GVLGNEDDGLHDDARDITFKLYFNKEYITSVBHYGKRLAAPNASASAMSPFTGY
MTMLKFTNRTTQVLSHSEYTEYEGTSFKLEEKDKHIVGYGKTESLQIGVYVVKP
IANA"
/join(18260..18398,18630..18859)
/note="gene id:MCL19.5"
/unknown protein"
/codon_start=1
/evidence=not_experimental
/protein id="BAB08246.1"
/db_xref="GI:9757721"
/translation="MSSSNKSNKSDHDDHNDQGHNTHTPSFMHTFETSNSSPS
SPVYSDPKPNQHOAQILEELFIGTWNPSLTSIKQITIKLQSYGEYVDDADVYK
WFNRKYSRKEPLYSYFMFL"
/complement(join(20751..20785,20889..20954,21047..21132,
21235..21269,21369..21455,21539..21616,21701..21741,
22239..22305))
/note="gene id:MCL19.6"
/unknown protein"
/codon_start=1
/evidence=not_experimental

/protein id="BAB08247.1"
/db_xref="GI:9757722"
/translation="MGRGCKFKGPTQORFFSAAADILAGTSAAPRSPKQKAEYBED
VEESEESESEEDADYKKGAEAVIEVDNPNVRQKTLKAKLDASKTTELSPRE
REELRQRAHERYMLQGGKTEQARKOLDRLALLRQREEAARKREKAARDAKKV
EGRK"
/complement(join(22934..23258,23634..23851))
/note="gene id:MCL19.7"
/pseudo
/codon_start=1
/evidence=not_experimental
/product="Rab-type small GTP-binding protein-like"
/join(27158..27278,27480..27606,27840..27915,28196..28338,
28394..28464,28552..28573,28658..28814)
/note="gene id:MCL19.8"
/pir||T27174
/similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein id="BAB08248.1"
/db_xref="GI:9757723"
/translation="MLGHSMDNYASDEVTSWAFLLLLAFSTRTVSPDCLEREER
KTMKRVRENEEKALWFRGQGNFDLVKCLILSKLNGHLHLSPPORDHRLLENL
NNHPLVLKQWERGQESLLRSEWISQSLVLSAITGLVSNASCNNHNLDMKH
LICKAACRICESFSTITALTEADIIYSEWIDECERVANTADDVVQEEEEEVEE
EHEEDEDHJVSVKRYNF"
/complement(join(29073..29934,30037..30602,30797..31014,
31113..31227))
/note="gene id:MCL19.9"
/codon_start=1
/evidence=not_experimental
/product="peptide transporter"
/protein id="BAB08249.1"
/db_xref="GI:9757724"
/translation="MTVEEVGDDYTDGTVDLRGNVRRSQTGRKACSFVVYVEVE
RMAYGISNNLVIYMTKLHQGVKSSNNVTNWSLTPILGAYVADAHPRGRIIF
VTSAYLGLMALLTSLVSLPGKPKPCSTANVENCKASVIGLVFFGALYLAIGT
GGPKNITIGADQDFEDPKDKIHKHSFNNWMSIFFGTFATVTVYQDNQWA
IGYCLTIGAFSIFLGLTRLYRHLKPMGSPPTKMARIVASLRKAREPMSDSRT
FYELPMEYAKKAPFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP
VLFVTFVSNMLAQIMTLFIKQGTDLDRLLTNFSPAPSLGFTTFESMLYSVLYDR
VYVFKRKLGTGNRGITLQRMGIMLILIMIIASITERLYKLVAAEGLHQTAV
PIPLSIFTLPOVLMGLADAFETIAKLEFFYDQAPESMKSGLSTSTYSMAVGFW
SIIILSSVQITKKQGRWIONNLNBSLDNYMFEAVNLNLFILVIRFVEYRAD
VTQSANVQKEPNVDNYNE"
/complement(join(32738..33188,33279..33677,33771..34336,
35792..36009,36633..36747))
/note="gene id:MCL19.10"
/codon_start=1
/evidence=not_experimental
/product="peptide transporter"
/protein id="BAB08250.1"
/db_xref="GI:9757725"
/translation="MTVEEVGDDYTDGTVDLQGNVRRSIRGRKACSFVVYVEVE
RMAYGISNNLVIYMTKLHQGVKSSNNVTNWSLTPILGAYVADALLGRIIF
VISCALYFSGMVLTVSTVPIGPKPCSTANVENCKASVIGLVFFGALYLAIGT
GGPKNITIGADQDFEDPKDKTOKLSFNNWMSIFFGTFATVTVYQDNQWA
LGVLPLGLAISITIFLLGTPFYRHLKPTGSPPTKMARIVASFRKANAPMTHDITS
FHELPSLEYERKGAFFLHPFSLRPLDRASLKTGTHHNLCTTTEVEETKQMLMLP
VLFITFVSNMLAQIMTLFVKQGTDLDRKVTGSPAPSLGFTTFESMLYSVLYDR
VYVFKRKLGTGNRGITLQRMGIMLILIMIIASITERLYKLVAAEGLHQTAV
KLPLTIFALLPOVLMGLADAFETIAKLEFFYDQAPESMKSGLSTSTYSMAVGFW
SIIILSSVQITKKQGRWIONNLNBSLDNYMFEAVNLNLFILVIRFVEYRAD
VTQSANVQKEPMKETE"
/join(39209..39382,39563..40135)
Alignment Scores:
Pred. No.: 1,5e-139 Length: 84510
Score: 1586.00 Matches: 323
Percent Similarity: 57.24% Conservative: 1
Best Local Similarity: 57.07% Mismatches: 1
Query Match: 90.58% Indels: 242
DB: 8 Gaps: 2

[illegible]

DB	QY	FT	Key	Location/Qualifiers	Location/Qualifiers	Length:	1518
Db	890	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	212	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	950	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	232	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	998	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	236	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	1058	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	255	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	1106	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	275	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	1133	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	295	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	1193	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	315	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	1253	QY	FT	Key	Location/Qualifiers	Length:	1518
RESULT 5							
BD012131							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES							
ORIGIN							
Alignment Scores:							
Pred. No.:							
Score:							
Percent Similarity:							
Best Local Similarity:							
Query Match:							
DB:							
US-09-787-737-2 (1-325) x E52178 (1-1518)							
QY	1	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	152	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	21	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	206	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	41	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	260	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	57	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	320	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	77	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	380	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	97	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	440	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	117	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	500	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	120	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	560	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	134	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	620	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	150	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	677	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	170	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	710	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	184	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	770	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	201	QY	FT	Key	Location/Qualifiers	Length:	1518
Db	830	QY	FT	Key	Location/Qualifiers	Length:	1518
QY	211	QY	FT	Key	Location/Qualifiers	Length:	1518

DB: 6 Gaps: 17

US-09-787-737-2 (1-325) x BD012131 (1-1518)

```

QY 1 MetSerSerAsnIlysAsnTrpProSerMetPheLysSerIysProCysAsnAsn 20
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 ATGGCTCTTCGAATAGACACTGGCCAAGCATGTTCAAGTCCAAACCT-----CATCC 205
QY 21 HisHisHisGlnHisGlnIleAsnThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 CATCAATGGCAACATGACATCACTCTCT-----CTCTGGCTTCTGCTTCTCACCGA 259
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProIysProArgTrp 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 TCTTCTCTCTTCTCTTCAGGATGTGAAGTGGAGAGGAGTCCAGAGCCAAACCAAGATGG 319
QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 AATCCAAAGCCAGACAGATTCGGATACTTGAAGCAATCTTTAACTCCGGGATGGTGAAT 379
QY 77 ProProArgGluGlnIleGlnArgIleArgLeuGlnGlnTyrGlyGlnIleGly 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 CCTCCAAGAGAGGAGATCAGAGGATAGGGCTCAGCTTCAAGATACAGGCCAATCGGT 439
QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 GATGCTAAAGCTTCTTACTGTGTTCCAAACCGTAAAGTCCCGTAGTAAACACAACTCCGC 499
QY 117 ValHisHis-----PheProPro----- 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 CTCCTCCCAACCACTCCAAACACTCTCTCCCTCAAAACCGAACCCGAGCGCGCAA 559
QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 560 CTTTCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 619
QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 620 AAAAGCAAGAAACAAGAACACTAAATCTCTTTGGGT---GGTAGTCAATGATCGGG 676
QY 150 LeuTyrProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPhePro 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 ATGTTTCCA-----CCGGAACCGGGG---TTTCTCTCTCCG 709
QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe--- 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 GTCTCCACTGTCGGAGGGTTGAAGGATATCACCGTCTCAUCCCAATTAGGGTTTCTCTCC 769
QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 770 GGTGATATGATTGAGCAACAAAAACCGCTCCAACGTGTACCGGACTCTCTCTGAGTGAG 829
QY 201 ValAsnAsnGly---ValAsnLeuGluThrAsnGlu----- 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 830 ATCATGAACGGTAGTGAGTTATGGAACATCATCAACACACTTGAAGTGAAGAAAGAA 889
QY 211 ----- 211
Db 890 GTTGAAGAAATGAGGATGAAGATGTTGCCAAGCCACAGACTCAGATTGTTAGCTTACC 949
QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyAsn 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 ACTAATCATCAATAGCTTCTTTACAAC-----AACCAACAACAACAACAATAAC 997
QY 232 Cys-----PheProPro----- 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 998 ATCATGCTTCATATTCTCTCCCACTACTCTACTGCCCACTATTACTACTTCGCATTCT 1057
QY 236 ---LeuThrValProLeuThrIleAsnGlnSerGlnGlnLysArgAspValGlyLeuSer 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1058 CTCGCTACTGTCCCATCACTCTCGGACCACTTCAAGTTCACGGGGAC----- 1105
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274

```

Db	1196	-----GCACGAATAAGAGTGTTCATCAATGAA	1192
QY	275	MetProfilGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla	294
Db	1193	ATGGAGCTTGAAGTCAGCTCAGGACCGTTCAATTGAGGGATGCATTTGGGGAGAGAGTT	1192
QY	295	ValleulleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThr-TyrGln	314
Db	1193	GTTCCTGATTAAATCCGGCGGTTCAGCCCATTCACCGCATGAATATGGCGTCTCTTCAC	1252
QY	315	ProLeuGlnAsnGlyAlaIleTyrTyrIleulle	325
Db	1253	CCTCTTCAACGACGAGCGCTCGTACTACTGATC	1285
RESULT 6			
LOCUS	AKI18501	1500 bp	linear
DEFINITION	Arabidopsis thaliana At2g3880 mRNA for putative homeodomain		PLN 06-DEC-2002
ACCESSION	AKI18501	transcription factor, complete cds, clone: RAFL19-73-K21.	
VERSION	AKI18501.1	GI:26452029	
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Arabidopsis thaliana full-length cDNA		
REFERENCE	Published Only in Database (2002)		
AUTHORS	2 (bases 1 to 1500)		
	Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:mseki@gsc.riken.go.jp, URL:http://pfweb.gsc.riken.go.jp, Tel:81-45-503-9625, Fax:81-45-503-9586)		
COMMENT	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.		

FEATURES
source

```

turner details.
location/Qualifiers
1..1500
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/Chromosome="2"
/clone="RAFL19-73-K21"
/note="common name: thale cress"
1..1500
/gene="At2g33880"
380..1297
/gene="At2g33880"
/codon_start=1
/product="putative homeodomain transcription factor"
/protein_id="EAC43105.1"
/db_xref="GI:26452030"
/translators="WVNPREDIRRAQLQBYGVQGDANFYFWQNRKSRSHKRLR
LNHSHXSPQTPQPPQPPASSSSSSSSSSKTKPRKSKNKNTNLSLGGSQMWG
MFPEPAFLPPVTSVGGFGFIYVSQLGSDMTQQKPAFTCTGLILSLIMGVS
YGHQOQLSKEKEVEEMRKMQLQPTQIYATNHOIASYNNNNNNNNIWLTPPTT

```

STATITTSHTSLATVPSTSDQLOVCADARIRVFINEMELELVSSGPFNVDRDAFGBEVL
INSAGQPIVTDYGVALPFLQHGASYILI"

ORIGIN

Alignment Scores:
Pred. No.: 8,94e-52 Length: 1500
Score: 643.50 Matches: 166
Percent Similarity: 50.61% Conservatives: 42
Best Local Similarity: 40.39% Mismatches: 84
Query Match: 36.75% Indels: 119
DB: 17 Gaps: 17

US-09-787-737-2 (1-325) x AK118501 (1-1500)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20
Db 161 ATGGCTTCTTCGAATAGACACTGGCCAGCATGTTCAAGTCCAAACCT-----CATCCC 214
QY 21 HisHisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
Db 215 CATCAATGACACATACATCACTCTCCT-----CTCTGGCTTCTGCTTCTCACCGA 268
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTrp 56
Db 269 TCTTCTCCCTTCTCTTCAGGATGTGAAGTGGAGAGGATCCAGAGCCAAACCAAGATGG 328
QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76
Db 329 AATCCAAAGCAGCAGCAGATTCGATACCTTGAAGCAATCTTAACTCCGGGATGTGAAT 388
QY 77 ProProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluIleGly 96
Db 389 CCTCCAAAGAGAGAGATCAGAGGATTTAGGGCTCAGCTTCAAGAAATACGCCAAGTCGGT 448
QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
Db 449 GATGCTAAAGCTCTTACTGTTCCAAAACCGTAAGTCCCGTAGTAAACACAACTCCGC 508
QY- 117 ValHisHis----- 119
Db 509 CTCCTCCACACCACTCCAAACACTCTCTCCCTCAAAACGCAACCGCAGCCGCGCAA 568
QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133
Db 569 CPTTCGGCTTCT 628
QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149
Db 629 AAAAGCAAGACAAAGACACACATAATCTCTCTTTGGT---GGTAGTCAAATGATGGG 685
QY 150 LeuTyrProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPhePro 169
Db 686 ATGTTTCCA-----CCGGAACCGGG-----TTTCTCTCCCG 718
QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183
Db 719 GTCTCCACTGTCGGAGGTTGAAGGTATCAGCTCTCAUCCCAATAGGTTTCTCTCC 778
QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200
Db 779 GGTATATGATTGACCAACAAAACCGGCTTCAACGTGTACCGGACTCCTCTGAGTGAG 838
QY 201 ValAsnAsnGly-----ValAsnLeuGluThrAsnGlu----- 211
Db 839 ATCATGAACGGTAGTGTGATTTATGGAATCATCATCAACCACTTGAAGTGAAGAA 898
QY 211 ----- 211
Db 899 GTTGAAGAAATCAGGATGAAGATGTGCAACAGCCACAGACTCAGATTGTTACGCTACC 958
QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyAsn 231
Db 959 ACTAATCATCAATAGTCTTCTTACAC-----AACCAACAAACAAACAAATAAC 1006

QY 232 Cys-----PheProPro----- 235
Db 1007 ATCATGCTTCATATTCCTCCCACTACTTCTACTGCCACCACTATTACTACTTCCGATTCT 1066
QY 236 ---LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
Db 1067 CTCGCTACTGTCCCATCACTCGGACCACTTCAAGTTCAAGCGGAC----- 1114
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274
Db 1115 -----GCAGGATAGAGATTTTCATCATGAA 1141
QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294
Db 1142 ATGGAGCTTGAAGTGAAGTCAAGCGTTCAGGATGTCATTTCCGGAAGAGTT 1201
QY 295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln 314
Db 1202 GTTCTGATTAATTCGCGGGTCAGCCCATTTGACCGGATGAATATGGCGTCTTTCAC 1261
QY 315 ProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
Db 1262 CCTCTTCAACAGGAGCCTCTGACTATCTGATC 1294
RESULT 7
U78721 114144 bp DNA linear PLN 27-FEB-2002
LOCUS Arabidopsis thaliana chromosome 2 clone T1B8 map TEN5, complete
DEFINITION sequence.
ACCESSION U78721 GI:20198308
VERSION U78721.3
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 114144)
Rounsley,S.D., Lin,X., Ketchum,K.A., Phillips,C.A., Brandon,R.C.,
Fuhrmann,J.L., White,O., Kerlavage,A.R., Adams,M.D.,
Somerville,W.R. and Venter,J.C.
Unpublished
REFERENCE 2 (bases 1 to 114144)
Lin.X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
JOURNAL 3 (bases 1 to 114144)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@igr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598813.
FEATURES
Location/Qualifiers
1. 114144
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="TEN5"
/clones="T1B8"
1. 1789
/note="overlap with BAC clone F4P9
(AC002332:150805..152593)."
1. 931
/gene="At2g33670"
/note="synonym: T1B8.33"
join(<1..153,246..281,358..425,534..594,692..>931)
/gene="At2g33670"
join(1..153,246..281,358..425,534..594,692..931)
/gene="At2g33670"

misc_feature

gene

mRNA

CDS


```

/note="Protein sequence is in conflict with the conceptual
translation"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69142.3"
/db_xref="GI:20198313"
/translation="KQAIISDWALEIOERHAVIQGMPLVNVSDRHFWRPALVLHI
IHILFQNAFEITYFFWYFGLRSCFHFFHALLIRVALGVGVQLCSYITPLVA
LVQMGSMTKRSVFDOTSKALKNWKNAKSKSTPGQTQPTPLNLRPKTGGLIESAS
PANITASVDVKESDQSQSRDLLSGP"
gene
2660..4843
/gene="At2g33680"
/note="synonym: TlB8.1; predicted by genscan and
genefinder, similar to GP 1402883 gnl PID e248504 X98130"
>2660..4843
/gene="At2g33680"
2660..4843
/gene="At2g33680"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69139.1"
/db_xref="GI:1707007"
/translation="MHPSTFQTELNPHSTLLKKLTHSQORNLVAGRAVHGQIIIRT
ASTCIGHANVLNPFYAKCKLAKASHISFNAILCKOVVSNLSLTGYSONGGISSYTV
MQFLREMRQDILPNAYTLAGIPKASSLSQSSQVGRQAAHVVKMSFGDIYVDTSLV
GMVYKAGLVDELGLKFAYPERNTYTWSTWVGATRGVBAIKVFNFLFREKERS
DSDYVFTAVLSLAATYVGLGRQIHCITIKNGLGFVALSNALVTMVKESLNEAC
KNPDSGDRNSITWSAMVTGYSONGSELEAVKLFSPMFAGIKPSYTIIVGLNACSD
ICVLEGRKOLHSLKLGPERHLPATLVDMTAKGLADARKDFDCIOERDVALWT
SLISGVQNSDNEBALIRRMKTAGIIPNDPTMASVLKACSLATLELGKQVHGHTI
KIGFGLVSPGALSMTYSKCSLEGNLFRTPNKDVVSNMAMISGLSHNGQDEA
LELFEMLAEGPEPDVPTVNIISASHKGFVRGWFYFNWMSDQIGLDPKDYACM
VLLSRAGOLKEAKETEGANIDHGLCLWRILLASCKNKGKCELGIVYAGEKLMALGR
ESSTYVQLSGITVLCGRMDREVRVWKMANGSVKSGVCSWELKNQIHFVVGDTMH
PMLEETKDLAVLSRQMIIEGFTVLDSSFVEEBEQTLSTFII"
5251..5466
/gene="At2g33690"
/note="synonym: TlB8.27; predicted by genscan and
genefinder"
>5251..5466
/gene="At2g33690"
5251..5466
/gene="At2g33690"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69143.1"
/db_xref="GI:2253578"
/translation="MSKSEKQELPETSPTKYEDIDYKKNAYGTSGHQDVKPGHG
GGTTDPTFGDAAPSAIDSANQAKK"
complement(6459..6492)
7252..9045
/gene="At2g33700"
/note="synonym: TlB8.2"
join(<7252..7434,7522..7713,7864..8069,8275..9045)
/gene="At2g33700"
join(7252..7434,7522..7713,7864..8069,8275..8836)
/gene="At2g33700"
/codon_start=1
/product="putative protein phosphatase 2C"
/protein_id="AAC69126.1"
/db_xref="GI:1707015"
/translation="WMSMDFSPLLTVLEGDNKNTSSATEIDTLENLDDTQISKGKP
PHLTSSATRLQLANADVDVNCVMSKLDKSPFLPYRSGCAEQAKQFMDEHI
CIDILVNLHGAIQCSSIGARYGYFDHGGTDAHFVKNILIRFIVDSDFPLCVKKA
IKSFLKADYEFADDSLIDSSITGATLTAFFIGRLITIANACDCAVLGRGRAIELS
KDHPCNTAEKREKLGGVYDGLNGQLSVARAIGDWMKPKGSACPLSPPELQ
ETDLSDEDFLIMGDCGLWVMSQCAVTIARKELMIHNDPERCSRELREALKNTIC
DNLTVIVCFPPDPQRIEIRMQSRVRSISAEGLNLLKGVLDGYP"
complement(11621..11649)
/rpt_family="AT rich"
repeat_region
11621..11649
/gene="At2g33710"
/note="synonym: TlB8.17; supported by full length cdna:
Ceres: 269036"
complement(join(<21321..21655,21744..21814,22051..22126,
22221..22298,22384..>22493))
/gene="At2g33735"
complement(join(21584..21655,21744..21814,22051..22126,
22221..22298,22384..>22446))
/gene="At2g33735"

```

[illegible]

Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones
Unpublished

2 (bases 1 to 918)

Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuai, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission
Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGRC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuai, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1. 918

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="2"

/clone="U60909"

/ecotype="Columbia"

/note="This clone is in pUNI 51"

1. 918

/note="putative homeodomain transcription factor"

/codon_start=1

/product="At2g33880"

/protein_id="AA063360.1"

/gb_xref="GI:28950873"

/translations="WNVPREETIRIRAQLQYGVGDANFYWFQNKRKSRSKHKLRL LNHSKHLPQTQPQPPQSASSSSSSSSSKSTKPKRKNKNTNLISLGGOMMG MPPEPAFLPVTVSGFGFTIVSQPLGSMDIEQQKPACTCTGLLSITMNGSVS WPGHQOHLSEKEVMKKMLQQOPTCYATTNHOIASYNNNNNNNIMLIPTT STATTITTSHTATVPSTDLOQVDARIVFINEMELEVSSGPFNVDFAGBEEVIL INSAGQPIVTDEYVALHPQGASVILI"

ORIGIN

Alignment Scores:

Fred. No.: 9.5e-31 Length: 918

Score: 420.50 Matches: 119

Percent Similarity: 44.74% Conservative: 30

Best Local Similarity: 35.74% Mismatches: 73

Query Match: Indels: 111

DB: 8 Gaps: 14

US-09-787-737-2 (1-325) x BT005296 (1-918)

QY 75 IleAsnProProArgGluGluIleLeGlnArgIleArgIleArgLeuGlnGlyrGlyGln 94

:::::::::::::::::::::
4 GTGAATCCTCCAAGAGAGAGATCAGGAGATTAGGCGTCACGTTCCAAGATA CGGCCAA 63

Dd

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

1 (bases 1 to 807)

Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 807)

Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

JOURNAL

COMMENT

Submitted (17-APR-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1..807
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U60620"
/ecotype="Columbia"
/note="This clone is in PUNI 51"

CDS

1..807
/note="unknown protein"
/codon_start=1
/product="At5g17810"
/protein_id="AAP12881.1"
/db_xref="GI:30017295"
/translation="MNQEGASHSPSTSTPEVARNWSPKPEQIILLISINSFTVNP
KDETVIRKMLEKFGVGDANFYVFNRRSRRRRRLQLAAATAATSGADHGH
MTAMSHQYPCSNNEIDIGFCSNLNGLSSSSQIPSPFGLGSSSSGCGENN
GMENLFMYGHSDHNNQHHSSNAASVLNPSDQNSQYEQEGFTVTFVNGPMEV
TKGALDKMTKFGDDSVLLHSSGLPLPTDFGFLMHLQHGQTYFLNPRQT"

ORIGIN

Alignment Scores: 2.45e-28 Length: 807
Pred. No.: 394.50 Matches: 105
Score: 47.93% Conservative: 38
Percent Similarity: 35.12% Mismatches: 101
Best Local Similarity:

Query Match:	22.53%	Indels:	55
DB:	8	Gaps:	9
US-09-787-737-2 (1-325) x BT006232 (1-807)			
QY	41	SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro	60
DB	19	TCACATAGCCCATCTCCACITCCACCGAACAGTCCGGGACGTTGGTCACTAACCG	78
QY	61	GlulnIleArgIleLeuSerIlePheAsnSerGlyThrIleAsnProArgGlu	80
DB	79	GAGCAAAATCTTCATCTACTCGAATCCATCTCAACAGTGGTACTGTTAACCCACAAAGAT	138
QY	81	GlulIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaSerVal	100
DB	139	GAACCGTGAGGTAAAGAAAGATGCTTGAGAAATTCGGTGTGGGAGACGCAACGTC	198
QY	101	PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys	120
DB	199	TTTCTACTGGTTTCAAAACCGACGCGTCAAGATCTCGC-----CGGAGACACCGCGCAG	249
QY	121	SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaHisCys	140
DB	250	CTTTTAGCAGCCACACCGACGCGCCACTCTCCATAGGAGCTGAAGACCCAGCACATG	309
QY	141	PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal	160
DB	310	AGCGCATGAGCATGATGATCA-----TATCCTTCGACGCAACCAACGAGATT-----	354
QY	161	ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerSerAlaGlnSerAla	180
DB	355	-----GAT	357
QY	181	PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr	200
DB	358	TTGGGGTTTGG-----AGTTGT	375
QY	201	ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle	219
DB	376	AGCACTTATCAGCTAATTAATCTTCTTAATGGATCGTCGTCATCTCAATCTCTCTTT	435
QY	220	AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys	232
DB	436	TTCTCTGGCGCTCTCTTCTTCAAGTGGTGGATGTGAGAACCAACAATGTTATGGAGAATCTC	495
QY	233	PheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysA-AspValGly	252
DB	496	TTCAAAATGATGGCCATGAATCTGATCATATCATCAGCAGCAGCATCATAGTCAAT	555
QY	253	-----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg---	267
DB	556	GCTGCATCAGTTTAAACCCATCTGATCAAACTCCCACTCCCAATACGAACAAGAGGG	615
QY	268	---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal	286
DB	616	TTTATGACCGTGTATTATAAAACGGGAGTTCTCTATGGAAGTAAACAAAGAGCAATAGACATG	675
QY	287	LysAlaAlaPheGlyAsnAspAlaValLeuLeuAsnSerPheGlyGlnProIleLeuThr	306
DB	676	AAACAATGTTTCGGTGAATGATTCGGTGTACTTCTTCTCTGCTCTCTCTCTCTCTCTCTCT	735
QY	307	AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle	325
DB	736	GATGAGTTTGGTTTCTTTCATGATCTTCTTACACATGACGACAACTTATTTCTCTGTA	792
RESULT 10			
AK117910			
LOCUS			
DEFINITION			
Arabidopsis thaliana At5g17810 mRNA for unknown protein, complete			
cds, clone: RAFL19-09-D18.			
ACCESSION			
AK117910			
VERSION			
AK117910.1 GI:26450881			
KEYWORDS			
FLI_CDNA; CAP trapper.			

Qy	94	nileglyAspAlaAsnValPheTyrTrpPheGlnAsnArglySerArgAlaIysHisLy	114
Db	113065	GGTCGGCGACGCGCAACGCTCTTCTACTGGTTCACAGACCGAAGTCCCGCACCAGAACAA	113124
Qy	114	sLeuArgVal-----HisHisLySer-----ProLysMetSe	125
Db	113125	GCTCGCGCGCGCGCGCCACCAACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGTC	113184
Qy	125	rLysLysAspLysThr-----ValIleProSerThrAspAla	137
Db	113185	GGCG	113227
RESULT 12			
AC121360		182868 bp DNA linear HTG 29-MAR-2003	
LOCUS		Oryza sativa (japonica cultivar-group) chromosome 5 clone	
DEFINITION		OJ1005 A08, ** SEQUENCING IN PROGRESS ***, 8 ordered pieces.	
ACCESSION		AC121360	
VERSION		AC121360.1 GI:20975769	
KEYWORDS		HTG; HTGS_PHASE2.	
SOURCE		Oryza sativa (japonica cultivar-group)	
ORGANISM		Oryza sativa (japonica cultivar-group)	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	
AUTHORS		1 (bases 1 to 182868) Chow, T.-Y., Hsing, Y.-I.-C., Chen, C.-S., Chen, H.-H., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R., Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Leu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y., Yu, S.-W., Wu, H.-P. and Shaw, J.-F.	
TITLE		Oryza sativa BAC OJ1005_A08 genomic sequence	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 182868) Chow, T.-Y. and Hsing, Y.-I.-C.	
AUTHORS		Direct Submission	
TITLE		Submitted (18-MAY-2002) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan	
JOURNAL		3 (bases 1 to 182868) Hsing, Y.-I.-C. and Chow, T.-Y.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (29-MAR-2003) Institute of Botany, Academia Sinica, 128, Section 2, Yen-chu-Yuan Road, Nankang, Taipei 11529, Taiwan	
TITLE		The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPC-Taiwan sequencing data.	
JOURNAL		* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.	
COMMENT		* This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.	
		1 12848 contig of 12847 bp in length	
		* 12948 contig of 12947 bp in length	
		* 12948 contig of 12947 bp in length	
		* 67635 contig of 54688 bp in length	
		* 67735 contig of 54688 bp in length	
		* 67736 contig of 17057 bp in length	
		* 67736 contig of 17057 bp in length	
		* 84793 contig of 17057 bp in length	
		* 84793 contig of 17057 bp in length	
		* 84793 contig of 17057 bp in length	
		* 105306 contig of 105306 bp in length	
		* 105306 contig of 105306 bp in length	
		* 105406 contig of 10994 bp in length	
		* 116400 contig of 116400 bp in length	
		* 116500 contig of 116500 bp in length	
		* 143095 contig of 143095 bp in length	
		* 143096 contig of 143096 bp in length	
		* 143196 contig of 143196 bp in length	
		* 153590 contig of 153590 bp in length	
		* 153590 contig of 153590 bp in length	
FEATURES		Location/Qualifiers	
source		1..182868	

Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T., Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshinara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. T., Eun, M. Y., Yano, M., Jiang, J. and Gojobori, T.

The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)

22373776
12447438

2 (bases 1 to 182681)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission
Submitted (13-JUN-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsukuba@ias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

On Apr 16, 2002 this sequence version replaced gi:16904697.
Genes were predicted from the integrated results of the following:
GENSCANL0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologs of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBST accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein- and '-like protein'. A gene covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21M13 to M13rev of the
BAC clone. This sequence of OSJNBb0063G05 clone has an overlap with
P0003808 (DBJ:AP03222) clone at the position 1 to 7853 of 5' end.
Detailed information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
Location/Qualifiers
1. 182681
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OSJNBb0063G05"
complement(join(292..470,2341..2400,2526..2714,2941..3038)
3164..3273,3366..3453,3563..3796))
/genes="OSJNBb0063G05.1"
complement(join(292..470,2341..2400,2526..2714,2941..3038)
3164..3273,3366..3453,3563..3796))
/genes="OSJNBb0063G05.1"
/note="contains ESTs AU032462 (S10523), D46091 (S10523)"
/codon_start=1
/product="putative phosphatidic acid phosphatase beta"
/protein_id="BAB90492.1"
/db_xref="GI:20161571"
translation="MPADARAAPSCPTTIRSHGAAVARSHAYDWLALLLVAVDGL
LNAIPFHRFVAGMDDIURPKMNTVPIWAVIATGPMIVFTVYFRERNVVDL
HVAIGILPSVLITGLTDIAKVGPRPREFWRCFPDGIATFVNVTGTVICHDGAS
VTRKDGKFPSPGTSWAGFLGSLVAGCTVFDPRGRHAKLCVILAPLLAVAWVA
ISRDVDDYHHMDVCTGGVLGVASVCYLOFEPAPSSCKWAPQWQSGCTTCTCTC

Mon Sep 13 10:30:57 2004

```
gene
SPRGEARTQSKPPNHWTPKPKVGEANEQAGSKTNLVNVIIRP"
join(8738..8870,10384..10573,11619..11676,13228..13308,
13401..13439)
/gene="OSJNBb0063G05.2"
join(8738..8870,10384..10573,11619..11676,13228..13308,
13401..13439)
/gene="OSJNBb0063G05.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90493.1"
/db_xref="GI:20161572"
/translation="MWIAHRHQHSCKSSPLATWVITLHKSSELSAARSDKLTFGS
SGGMKTSOTAPKTLVPLPPKPPQPPKPPKQGLVDRRRPKPPT
AFFYPMEDFRKTYKEENPSVKMQEERVKYDIATKRAEYKVAEAFDKKKSELS
RESVDY"
complement(join(14070..15105,16146..16285))
/gene="OSJNBb0063G05.3"
complement(join(14070..15105,16146..16285))
/gene="OSJNBb0063G05.3"
/note="hypothetical protein
similar to Oryza sativa chromosome 1, P0046E05.21"
/codon_start=1
/protein_id="BAB90494.1"
/db_xref="GI:20161573"
/translation="WAKTAPEQKAVTPATPAASAPFFHFVYVGNLSSTTWRDLR
SSWKANVYRMVIACTIOGVYLLEDRQERRDERTGALAPQWRPFKYKLAQALVDER
DGSVYATQWDRQALCDYIPPRPAGAAVVALRGITLLRPAFTFRDVTDLIRFLAW
DSLGKVRPAGALAAARATGRGAGSCVGHSLGALGTFALQVKGALAKGVFVECHV
FNPPSVLSAMSLRGFAETAGELWGRVSWIPYGGSSSSQAADAGGGESEFAKAM
RPHLYININDIICYYNDAAGATATVAGGGGGGSGAKVAAAAAVRGTTGGGA
RVARMLVASGPTKFFLEAGLEQWADDVLEQVALNHSKLIDRLSLAAAPPAAS"
complement(join(17436..17616,19365..19735))
/gene="OSJNBb0063G05.4"
complement(join(17436..17616,19365..19735))
/gene="OSJNBb0063G05.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90495.1"
/db_xref="GI:20161574"
/translation="MSPSGSGPSASAKDTDAQAQOEQPFQKGGATRRERMESLW
LTSAVMPKKGALIEGVGAASILLDLKAQVLTQBEARKPTAHDAASAAAASGEFRRAK
KRAAGDPGAKNSGVDAARAKHVEHSEARQKASTMKSRQEQEAAREKLQAYL
KKLEKLELAEKRAISATDDQPAS"
join(20120..20306,21024..21089,21152..21495)
/gene="OSJNBb0063G05.5"
join(20120..20306,21024..21089,21152..21495)
/gene="OSJNBb0063G05.5"
/note="contains EST (S14119)
unknown protein"
/codon_start=1
/protein_id="BAB90496.1"
/db_xref="GI:20161575"
/translation="WACSFPPSSATRLQAVDAAVRSRVPRAATVSPVQRSIAGC
KAAARQEGATQPAISVSSARTQLDLLEOLTPTSDGIAYVATNTHSSGVENGAP
TEPRVQRTIRQLSAIGDRDGETILBLGKLLKSLNLSVQSRNIKRQALLTK
VSGRNDVFFATVGAFVLPVPLAILAVALTGYVQLLP"
complement(join(22226..22535,22731..22794,23018..23029,
24891..25001,27666..27752,27942..28086,29521..29655))
/gene="OSJNBb0063G05.6"
complement(join(22226..22535,22731..22794,23018..23029,
24891..25001,27666..27752,27942..28086,29521..29655))
/gene="OSJNBb0063G05.6"
/note="contains EST AU032335 (R3907)
unknown protein"
/codon_start=1
/protein_id="BAB90497.1"
/db_xref="GI:20161576"
/translation="WVELVWVVTTEIVFEERODERBWTKEEDWEKKNDKNYESWTWM
QASGNRLVATASASLLQTSAGAFIRFKSSGCGTHIKLFEDMQIKRPFCEYIOGL
TNMKFOOIKLTTEFADSCPSRRGVKRSVLSWFLHWLGLTATGATGNTVYGLHT
DTQKHQIIICLTTEFADSCPSRRGVKRSVLSWFLHWLGLTATGATGNTVYGLHT
YHRTTKSVKLTWGLTTFELSLLLFFYLLIDRWSYMMKQGNAPIEQLRTDNRKTYPT
TLRKELGWQVE"
gene
join(31079..32110,32689..32979,34323..34389,35213..35331,
35415..35507,35915..36004,36920..37042,37225..37298,
37395..37449,37604..37839,38476..38605,38687..38789,
38945..39042,39769..39936)
/gene="OSJNBb0063G05.7"
join(31079..32110,32689..32979,34323..34389,35213..35331,
35415..35507,35915..36004,36920..37042,37225..37298,
37395..37449,37604..37839,38476..38605,38687..38789,
38945..39042,39769..39936)
/gene="OSJNBb0063G05.7"
/note="contains ESTs AU032273 (R3833), AU032274 (R3833)"
/codon_start=1
/product="putative vacuolar protein sorting protein"
/protein_id="BAB90498.1"
/db_xref="GI:20161577"
/translation="MGYRVELVPIRLDPVHGLTVPSRSPROSTSLTAHLARDAAAA
SPPMSSSVSVAEWDLLSDSEFYRITVYSLPWPSPPSAAAASSSTSGGGGGGSG
GVLGRDLSTHIVAAAPFGGFIAAVRDSDSKIVQLHSEPSRRELLLYSSGHPIASSPW
PPQLPRLSLAFSSSLNVLALLSGSLRFLPDKNPSPTPVLPTSSGSGVADAA
FWGGVAVLTEDNRVVVTTDADDPHREFADPCVQDEQVLCMAVPEPQVMSGSP
EVLLAVGRVLAVDDDDVTGLGLELEIGVQKMAVSPNGKLLAAFAHDRLVLPDF
SKLIFYECDSALPDQIAWGLDVLVYMSAALMVGNGDPVLYNDPEIKLPIEC
DOVRILSNMFEFLHRVPDSTISIFIGMSPAALYDARDHYDQSAKIPHKPIPT
DUFIKIQAARVILVGFQSPHERPOEMCKTLVNAVDPQIGMPLTQOQLLKP
VLIGLIANAOHLALRISEYLNPEVIMHWACEKITASAIDPTVLLEGLLKIP
LLLSIDQDKALSKAIESGDTDLVYLVLVHFWKVAVERKAPDPGVINARPLARDL
LWYARHSHKALKDFLSTGLQVLSFMSFYFLNREKIQADAPFLLLKSELEARN
PNASGPHLGPQVRLIEQARHLFAETKEHVESKASEEHLKSHQHEVSTKQAI
FVSGSVSDITKTCIANGNERAALKVKEFKPKWYWLKSCALATVGNWMDALEKFS
ERRPGGYKPFVEACIDAGQKTEALKYIKPLTPRSEAYARIKMAKEAFAAASQVK
DSEDLFRLKLTLAQNTAAASIFDLRLDSFQGT"
complement(join(40616..40957,41753..41867,41920..42035,
42111..42125))
/gene="OSJNBb0063G05.8"
complement(join(40616..40957,41753..41867,41920..42035,
42111..42125))
/gene="OSJNBb0063G05.8"
/note="contains EST AU032380 (R3969)"
/codon_start=1
/product="putative 60S ribosomal protein L18a"
/protein_id="BAB90499.1"
Alignment Scores:
Pred. No.: 8,04e-21 Length: 182681
Score: 347,50 Matches: 77
Percent Similarity: 52.78% Conservative: 18
Best Local Similarity: 42.78% Mismatches: 21
Query Match: 19.85% Indels: 65
DB: 3 Gaps: 3
US-09-787-737-2 (1-325) x AP003760 (1-182681)
Qy 1 MetSerSerAanLysAasnTrrProSerMetPhelySerLysProCysAasnAsn 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66212 ATGCGTCGTGCAAGCAGCACTGGCCGAGCATGTTTCAGTGAAGCAGCAGCAGCG 66153
Qy 21 HisHisHisGlnHisGluile-----AspThrProSerTyrMetHisTyrSerAsnCys 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66152 TGGCAGCAGCAGCTGACATGCGCGGCTCCCTCCTCTCGGCTCCTCGCC 66093
Db 39 Asn-----LeuSerSerSerPheSerSerAsp-----47
Qy 66092 GGCAGCGCGCGGGCGGCGGCTACTCCCTCAAGTCGTCGCCCTTCTCGT 66034
Db 47 -----47
Qy 66033 GATCGATCTCGCGCGCGCATCCGTCGCTACTGTTTCGGCTTTTGTACTATG 65974
Qy 47 -----47
65973 AGGGTTTCGTGTTGCGAGGTTACAGGCGATCGAGATCGATCGGAGGGAATTATGG 65914
Db 48 -----ArgIleProAspProLysProArgTrrp 56
Qy
```



```

Db      65913  GATTTCGGTGTGTTGCTTGCAGTGGCGAGGAGGTTCCGGACCCGGAAGCGCGTGG 65984
QY      57  AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76
Db      65953  AACCCGGCGCGGAGCAGATCCGATCTCGAGGCGATCTTCAATCCGGGATCGTCAAC 65794
QY      77  ProProArgGluGluIleGlnArgIleArgIleLeuGlnGluTyrGlyGlnIleGly 96
Db      65793  CCGCGCGCGGAGCAGATCCCGCGCATCCGATCGAGCTGCGAGGATCGGCGCAGTCCGC 65734
QY      97  AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisGlyLeuArg 116
Db      65733  GACCCCAACGCTTCTACTGTTCTTCAGAACCGCAAGTCCCGCTCCAGAACAGTCCGC 65674

RESULT 14
AK063503
LOCUS   Oryza sativa (japonica cultivar-group) cDNA clone:001-116-B11, full
DEFINITION
insert sequence.
ACCESSION
AK063503
VERSION 1.1
KEYWORDS
FULL CDNA; oligo-capping.
SOURCE  Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
1 The Rice Full-Length cDNA Consortium, National Institute of
  Agricultural Sciences Rice Full-length cDNA Project Team:
  Kikuchi,S., Satoh,K., Negata,T., Kawagashira,N., Doi,K.,
  Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Hotta,I.,
  Kojima,K., Namiiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
  Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
  Science Genome Sequencing & Analysis Group: Ohtomo,Y., Murakami,K.,
  Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
  Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
  Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
  Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
  Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Yoshimura,A., Miura,J.,
  Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
  Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
  Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
  Sato,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Yoshino,M., and Hayashizaki,Y.
  Collection, mapping, and annotation of over 28,000 cDNA clones from
  japonica rice
  Science 301 (5631), 376-379 (2003)
  2752273
  12869764
  2 (bases 1 to 1282)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

```

```

TITLE
JOURNAL
COMMENT
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
  Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
  Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiiki,T.,
  Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., and
  Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ohtomo,Y., Iida,Y.,
  Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
  Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
  Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
  Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
  Yoshimura,A., Matsubara,K., and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
  Akimura,T., Arakawa,T., Hayashida,K., Hayatsu,N., Hiramoto,K.,
  Hara,A., Hashidume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
  Hiraoka,T., Hori,F., Iida,Y., Imamura,K., Imotani,K., Ishii,Y.,
  Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kondo,S.,
  Kishikawa-Hirozane,T., Kojima,Y., Konno,H., Kondo,H., Kouda,M.,
  Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
  Nakamura,M., Nishi,K., Nomura,K., Numasaka,R., Ohno,M., Osato,N.,
  Ota,Y., Saitoh,H., Sakai,K., Sakai,K., Sakazume,N., Sano,H.,
  Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
  Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
  Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
  Yasunishi,A., and Hayashizaki,Y.
Location/Qualifiers
1..1282
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="cDNA"
/db_xref="taxon:39947"
/clone="001-116-B11"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1.11e-22 Length: 1282
Score: 338.00 Matches: 89
Percent Similarity: 46.34% Conservative: 44
Best Local Similarity: 21.01% Mismatches: 92
Query Match: 19.30% Indels: 62
DB: 8 Gaps: 9

```

```

US-09-787-737-2 (1-325) x AK063503 (1-1282)
QY 53 LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer 72
Db 171 AGGTCGGCGTGGAGCGCGAAGCGGAGCAGATCTCATCTCTGGAGTCCATCTTCAACAGC 230
QY 73 GlyThrIleAsnProProArgGluGluIleGlnArgIleArgLeuGlnGluTyr 92
Db 231 GGCATGCTGTAACCGCCGACGAGCAGACCGTCCGCGATCCGCAAGCTGCTGAGCGCTTC 290
QY 93 GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys 112
Db 291 GGGCGCCGCGGAGCGCAACGCTTCTACTGGTTCAGAACCGCCGCTCGCTCCGCG 350
QY 113 HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle 132
Db 351 CGCCGCGCAGCGC----- 362
QY 133 ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro 152
Db 363 ---CAGTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419

```

```

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
  Agricultural Sciences, Department of Molecular Genetics, Head of
  Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki
  305-8602 Japan (E-mail:skikuchi@nias.affrc.go.jp,
  Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
  rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/

```

Qy	153	ValGlnAenAenGluLeuValThrGluProAlaGlyPheLeuPheProValHis---	171
Dd	420	-----ACTGCCTCGCGTGGTGCCGTTCGCCACAGCC	455
Qy	172	AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProValVal	191
Dd	456	GCTCGCGCGGTTCGCTCGCTGGATGTTCCGCCACAGC	494
Qy	192	ThrGluGluGlyMetAlaPheSerThrValAsnAenGlyValAenLeuGluThrAsnGlu	211
Dd	495	----GCGCGCGGTTCAGCTCTCGCTCGCTTCATCGCTCGCGCTCC	536
Qy	212	AsnPheAspLysIleProAlaIleAsnLeuTyrglyglyasp-----GlyAsnGly	228
Dd	537	-----TCGCGCGCGGTTCGCTGGATGATGATGCGGACCTGGACTACGCGCGCGC	587
Qy	229	GlyGlyAsnCysPheProLeuThrValProLeuThrIleAenGlnSerGlnGluLys	248
Dd	588	GCGCACGACTGTT-----GCCATCTCGAGGACAGATGGGTAC	626
Qy	249	ArgAspValgly-----LeuSerGlyGlyGluAspValGlyAspAsn	262
Dd	627	ATGCACGCGCGCGCGCTCGCTCGTGGCGCGCGCGGTCACATCATCAGCAGCAGCAG	686
Qy	263	Valtyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu	278
Dd	687	CTCTACTACTCGTGTCACTCGCAGCATGCGTGTTCATCAACGAGTGGCGCAGCGAG	746
Qy	279	ValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsn	298
Dd	747	GTGCCAAGGCGCGCATCGATCGATCAATGTTTGGCGCAGCATGATGCTGTGTGAT	806
Qy	299	SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrglnProLeuGlnAsn	318
Dd	807	TCAAAGGCTGCTCTTCTTCACGACGAGTAGCAGCATCTCTCTCCATTCTCTCCAGATG	866
Qy	319	GlyAlaIleTyrrTyrrLeulle 325	
Dd	867	GCGCAGAGTACTTTCCTGGTTC 887	
RESULT 15			
AK073232			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J03302B15, full insert sequence.		
ACCESSION	AK073232		
VERSION	AK073232.1 GI:32983255		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.		
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team:		
	Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yarakai, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kishimoto, N., Nishi, K., Imamura, K., Imotani, K., Li, C., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Miura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Matsubara, K., RIKEN: Kusumegi, T., Oka, M., Ryū, R., Ueda, M., Kobayashi, M., Xie, Q., Lu, M., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.		
JAPONICA rice			
TITLE			

Science 301 (5631), 376-379 (2003)
 22752273
 12869764
 2 (bases 1 to 1378)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imoto, K., Kishimoto, N., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kikuchi, S., Kodama, T., Kodama, T., Kuroki, Y., Kudo, S., Konno, H., Kouda, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kudo, M., Koyama, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Ohtsuki, K., Oka, M., Osato, N., Ota, Y., Otomo, Y., Ryū, R., Saitoh, K., Satoh, K., Shibata, K., Sakabe, Y., Sano, H., Sasaki, D., Sato, K., Shihashi, T., Sugabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyama, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skkichi@nias.affrc.go.jp]
 Tel:81-29-838-7007, Fax:81-29-838-7007
 This clone is one of the 28K full-length cDNA clones from japonica rice.
 URL : http://cdna01.dna.affrc.go.jp/cdna/
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohtsuki, K., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.
 FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryū, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
 Location/Qualifiers
 1. 1378
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J03302B15"

ORIGIN			
Alignment Scores:			
Pred. No.:	1,22e-22	Length:	1378
Score:	338.00	Matches:	89
Percent Similarity:	46.34%	Conservative:	44
Best Local Similarity:	31.01%	Mismatches:	92

Query Match:	19..30%	Indels:	62
DB:	8	Gaps:	9
US-09-787-737-2 (1-325) x AK073232 (1-1378)			
QY	53	LysProArgTTPAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer	72
Db	163	AGGTGCGGTGTGACGCGCGAAGCGGAGCAGATACTCATCTGGAGTCCATCTTCACAGC	222
QY	73	GlyThrIleAsnProProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyr	92
Db	223	GGCATGGTGAACCCGCCCAAGACAGACAGCCGTCGACATCCGAAGCTGCTCGAGCGCTTC	282
QY	93	GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys	112
Db	283	GGCGCGGTGCGGACGCGCAACGCTTCTTACTGTTCCAGAACCGCGCTCGCGCTCCGCG	342
QY	113	HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle	132
Db	343	CGCCGCCAGCGC-----	354
QY	133	ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro	152
Db	355	---CAGCTGCGAGCGCACGGCGAGCGCGCGCGCGCGCTCGTCGGGATCTCTCCG	411
QY	153	ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis---	171
Db	412	-----ACTGCTCGTCGGTGCCTCGCGCTCGCGCTCGGCACGCC	447
QY	172	AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValValProValVal	191
Db	448	GGCTCGCGGGTCTCGTCTCGGGATGTTCCGCGCAGGC-----	486
QY	192	ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu	211
Db	487	---GCCCGCGGGTACAGCTCTCGTCTCTCATCTGCGCGCTC-----	528
QY	212	AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly	228
Db	529	-----TCGCCGCGCTCGTGGGATGATGATGGGGACGTGTGACTACGGGGCGGC	579
QY	229	GlyGlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys	248
Db	580	GGCGACGACCTGTTC-----GCCATCTCGAGCGAGATGGGGTAC	618
QY	249	ArgAspValGly-----LeuSerGlyGlyGluAspValGlyAspAsn	262
Db	619	ATGACGCGCGCGCGCTCGTCTGTCGCGCGCGCGCGTCCAGATCAGCAGCAGCAG	678
QY	263	ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu	278
Db	679	CTCTACTACTGCTGTCACACCTTCGACGATACACGGTGTTCATCAACGAGTGGCGAG	738
QY	279	ValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsn	298
Db	739	GTSCCAAGGGACCGATCGATCTGAGATCAATGTTTGGGCGAGGACGTGATGCTGTGCAT	798
QY	299	SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsn	318
Db	799	TCAACGGGTGCTTCTTCCAGCCACGAGTACGGCATCTCTCTCCATTCCTCCAGATG	858
QY	319	GlyAlaIleTyrTyrLeuIle	325
Db	859	GGCGAGAGCTACTTCTCTGTC	879

Search completed: September 10, 2004, 10:47:34
Job time : 4872 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:25:41 ; Search time 471 Seconds
(without alignments)
2931.347 Million cell updates/sec

Title: US-09-787-737-2
Perfect score: 1751
Sequence: 1 MSSSNKNWPMFKSKPCNNN.....TDFGVGYQLQNGAIYYLI 325

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q=/cgn2.1/USPTO spool h/US0978737/runat 01092004 112454 23588/app query.fasta 1.519
-DB=N Geneseq 29Jan04 -QWNT=fastap -SUPFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-NODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0978737@cgn 1.1 470 @runat 01092004 112454 23588 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	% Match	Query Length	ID	Description
1	1751	100.0	1214	5	Aaf59786 Arabidops
2	645.5	36.9	1518	5	Aaf59787 Arabidops
3	623	35.6	1502	9	Add30297 Plant yie
4	394.5	22.5	807	6	Abk65270 Arabidops
5	394.5	22.5	807	9	Add30133 Plant yie
6	342	19.5	994	6	Abk65367 Arabidops
7	328.5	18.8	1367	4	Aas02222 Corn Wusc
8	212.5	12.1	816	9	Add30299 Plant yie

9	212.5	12.1	1326	3	AAC38580 Arabidops
10	200	11.4	1209	3	AAC47056 Arabidops
11	189.5	11.4	1227	9	Add30291 Plant yie
12	188	10.7	371	3	AAC56282 Arabidops
13	187.5	10.7	1338	4	AAS02219 Corn Wusc
14	187	10.7	741	4	AAS02227 Arabidops
15	175.5	10.0	844	4	AAS02226 Arabidops
16	173	9.9	904	4	AAS02229 Arabidops
17	172.5	9.9	906	4	AAS02228 Arabidops
18	171.5	9.8	522	7	AAD47508 Arabidops
19	171.5	9.8	648	3	AAC55917 Arabidops
20	171.5	9.8	717	6	ABQ65592 Arabidops
c	170	9.7	506	4	AAS02225 Soybean W
21	170	9.7	1190	2	AAS02289 Arabidops
22	170	9.7	720	4	AAS02220 Corn Wusc
23	167.5	9.6	720	4	ABK65267 Arabidops
24	167.5	9.6	783	6	ABK65267 Arabidops
25	167.5	9.6	783	9	ADD30617 Plant yie
26	167.5	9.6	783	9	ADE31512 Arabidops
27	167.5	9.6	1136	2	AAS20288 Arabidops
28	167	9.5	945	3	AAC42316 Arabidops
29	166	9.5	767	4	AAS02221 Corn Wusc
30	140.5	8.0	1194	4	AAS02224 Arabidops
31	135.5	7.7	513	4	AAS02223 Arabidops
32	129	7.4	441	3	AAC55958 Arabidops
33	117	6.7	4791	8	ADA30129 Arabidops
34	113	6.5	771	6	ABS63002 Arabidops
35	113	6.5	771	6	ABT11607 Arabidops
36	112.5	6.4	1302	4	AAD06478 Arabidops
37	112.5	6.4	1302	9	AD37278 Plant yie
38	111.5	6.4	7377	7	ACF69433 Arabidops
39	111.5	6.4	110000	7	ACF67367_23 Arabidops
c	111.5	6.4	110000	7	ACF65386_5 Arabidops
41	111	6.3	2287	6	ABA95489 Arabidops
42	109.5	6.3	462	4	AAS02218 Arabidops
43	108.5	6.2	3082	9	ADB62521 Arabidops
44	108.5	6.2	3845	7	ACC72452 Arabidops
45	108.5	6.2	3886	6	ABQ75807 Arabidops

ALIGNMENTS

RESULT 1	
ID	Aaf59786 standard; cDNA; 1214 BP.
XX	XX
AC	Aaf59786;
XX	XX
DT	04-MAY-2001 (first entry)
XX	XX
DE	Arabidopsis thaliana M6 homeodomain protein cDNA.
XX	XX
KW	Homeodomain-like; homeobox gene; M6; plant differentiation; MSH;
KW	many shoot; expression control; adventitious bud formation;
KW	branching induction; cytokinin independent; transgenic plant; crop;
KW	agriculture; ss.
OS	Arabidopsis thaliana.
XX	XX
PN	W0200107618-A1.
XX	XX
PD	01-FEB-2001.
XX	XX
PF	21-JUL-2000; 2000WO-JP004904.
XX	XX
PR	22-JUL-1999; 99JP-00207995.
XX	XX
PA	(SUNR) SUNTORY LTD.
XX	XX
PI	Kakimoto T;
XX	XX
DR	WPI; 2001-182796/18.
DR	P-PSDB; AAB60642.
XX	XX

PT Arabidopsis thaliana originated homeobox genes encoding proteins
 PT participating in differentiation with ability to promote adventitious bud
 PT formation and branching induction, applicable in improving plant cells
 PT and plants e.g. for crops.
 XX

XX Claim 3; Page 20-23; 34pp; Japanese.

PS The invention relates to two novel Arabidopsis thaliana proteins M6 and
 CC M8 (AA60642 and AA60643, respectively), and to the cDNAs encoding them
 CC (AA60642 and AA60643). The proteins contain homeodomain- like sequences
 CC and participate in differentiation, having the ability to promote
 CC adventitious bud formation and branching induction by controlling
 CC expression of the many shoot (MSH) gene without the requirement for
 CC cytokinin. The invention also relates to expression vectors and host
 CC cells which contain M6 or M8 nucleic acid sequences, the recombinant
 CC production of the M6 or M8 protein, a plant or plant cell transfected
 CC with the M6 or M8 gene, and methods for inducing differentiation,
 CC adventitious bud formation and/or plant branching via expression of the
 CC transacted M6 or M8 gene. Nucleic acids encoding the homeodomain
 CC proteins may be used to generate transgenic plants with improved
 CC adventitious bud formation and branching for use e.g., as agricultural
 CC crops. The present sequence represents cDNA encoding the Arabidopsis
 CC thaliana M6 protein
 XX

SQ Sequence 1214 BP; 395 A; 228 C; 226 G; 365 T; 0 U; 0 Other;

Alignment Scores: 1214
 Pred. No.: 1,13e-180 Length: 1214
 Score: 1751.00 Matches: 325
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB:

US-09-787-737-2 (1-325) x AAF59786 (1-1214)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20
 DB* 36 ATCTCTCTCTCAACAAATAATGGCCAAAGCATGTTCAAAATCCAAACCTTGCAACAATAAT 95
 QY 21 HisHisGlnHisGlnLysAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
 DB 96 CATCATCATCAACATGAATTCGATACCTCATCTTACATGCACTACTCTATTTCACACTA 155
 QY 41 SerSerSerPheSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60
 DB 156 TCATCTCTCTTCTCAGATCGATACGATACGATCTTAAACCGAGATGGAATCTCTAAACCG 215
 QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80
 DB 216 GAGCAGATTAGGATACCTGAATCAATCTTCAATTCGGTACTATTAAACCCACCTAGAGAG 275
 QY 81 GluIleGlnArgIleArgLeuGlnGlnLysGlyGlnIleGlyAspAlaAsnVal 100
 DB 276 GAGATTCAAGAAATCCCGATCCGCTTCAAGAAATGTTCAATTCGATCGAGCGCAACCGTG 335
 QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120
 DB 336 TTTTACTGTTTCAAAACCGGAATCTCGAGCAAAACCAAGCTTCGTGTTTCATCACAAA 395
 QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
 DB 396 AGCCCTAAATGTCAAGAAAGACAAAGACGCTTATCTTAGTACTGACGCTGATCATTTGT 455
 QY 141 PheGlyPheValAsnGlnGlnLysGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160
 DB 456 TTGTGTTTGTAAACCAAGAACCGGATATATCCGGTTCAAAACAATGAGTTGTGTGTA 515
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
 DB 516 ACCGAAACCGCGCGTTTCTATTTCGGTTTCATATGATGATCCGAGCGCTCTCAATCAGCG 575
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200

Db 576 TTGGTTTTCGCGATTTTGTGTACCGGTGTTACCGAAGAGGATGGCATTTCTTACC 635
 QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsn 220
 Db 636 GTTAATAACGCGTAAATTTGGAGACTAACGAAAATTTTGATAAAATTTCCGCGCATCAAT 695
 QY 221 LeuTyrGlyAspGlyAsnGlyValGlyAsnCysPheProProLeuThrValProLeu 240
 Db 696 TTATACGGCGAGATGGAAATGGCGGTGGAAATGTTTCTCTTCTGACTGTTCCATTA 755
 QY 241 ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGly 260
 Db 756 ACCATCATCAATCTCAAGAAAAACGAGATGTAGGATTTATCCGGTGGTGAAGACGTCCGA 815
 QY 261 AspIleValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValVal 280
 Db 816 GATATGTTTATCCGGTGAGATGACGGTGTATTATACGAGATGCCCTATCGAAGTAGTG 875
 QY 281 SerGlyLeuPheAsnValIysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPhe 300
 Db 876 TCTGATTTATTCACGTTTAAGCGACGTTTCGAAAACGATGCCGTTTGTATCAACTCGTTT 935
 QY 301 GlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAla 320
 Db 936 GGCCAGCGCTATTCTTACAGATGATTTGGTGTATTCTATCAACCTCTCCAAAATGGCGCA 995
 QY 321 IleTyrTyrLeuIle 325
 Db 996 ATCTATTATCTTATT 1010
 RESULT 2
 AAF59787
 ID AAF59787 standard; cDNA; 1518 BP.
 AC AAF59787;
 XX
 XX 04-MAY-2001 (first entry)
 DT
 DT Arabidopsis thaliana M8 homeodomain protein cDNA.
 DE
 XX Homeodomain-like; homeobox gene; M8; plant differentiation; MSH;
 KW many shoot; expression control; adventitious bud formation;
 KW branching induction; cytokinin independent; transgenic plant; crop;
 KW agriculture; ss.
 KW
 XX Arabidopsis thaliana.
 OS
 XX
 XX WO200107618-A1.
 PN
 XX
 XX 01-FEB-2001.
 PD
 XX 21-JUL-2000; 2000WO-JP004904.
 PF
 XX 22-JUL-1999; 99JP-00207995.
 PR
 XX (SUNR) SUNTORY LTD.
 PA
 XX Kakimoto T;
 PI
 XX WPI; 2001-182796/18.
 DR
 XX P-PSDB; AAB60643.
 DR
 XX Arabidopsis thaliana originated homeobox genes encoding proteins
 PT participating in differentiation with ability to promote adventitious bud
 PT formation and branching induction, applicable in improving plant cells
 PT and plants e.g. for crops.
 XX
 PS Claim 6; Page 25-28; 34pp; Japanese.
 CC The invention relates to two novel Arabidopsis thaliana proteins M6 and
 CC M8 (AA60642 and AA60643, respectively), and to the cDNAs encoding them
 CC (AA60642 and AA60643). The proteins contain homeodomain- like sequences
 CC

and participate in differentiation, having the ability to promote adventitious bud formation and branching induction by controlling expression of the many shoot (MSH) gene without the requirement for cytokinin. The invention also relates to expression vectors and host cells which contain M6 or M8 nucleic acid sequences, the recombinant production of the M6 or M8 protein, a plant or plant cell transfected with the M6 or M8 gene, and methods for inducing differentiation, adventitious bud formation and/or plant branching via expression of the transfected M6 or M8 gene. Nucleic acids encoding the homeodomain proteins may be used to generate transgenic plants with improved adventitious bud formation and branching for use e.g., as agricultural crops. The present sequence represents cDNA encoding the Arabidopsis thaliana M8 protein.

SQ Sequence 1518 BP; 425 A; 389 C; 279 G; 425 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	4,56e-60
Score:	645.50
Percent Similarity:	50.61%
Pest Local Similarity:	40.39%
Query Match:	36.86%
DB:	5
	Gaps: 17
	Indels: 119
	Mismatches: 84
	Conservative: 42
	Matches: 156
	Length: 1518

US-09-787-737-2 (1-325) x AAF59787 (1-1518)

[illegible]

QY	201	ValAsnAsnGly---	ValAsnLeuGluThrAsnGlu	-----	211
		:::		:::	
Db	830	ATCATGAACGGTAGTGTTGAGTTATCGAACTCATCATCAACACTGTGAGTGAGAAAGAA	889		
QY	211	-----	-----	-----	211
Db	890	GTTGAAGAAATGAGATGAAGATGTTGCAACAGCCACAGACTCAGATTTGTTACGCTACC	949		
QY	212	AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyGlyAsn	231		
		:::		:::	
Db	950	ACTAATCATCAATCAAGTCTTCTCAAC-----AAACAACAACAACAATAAAC	997		
QY	232	Cys-----PheProPro-----	235		
Db	998	ATCATGCTTCATATCTCTCCCACTACTTCTACTGCCACCACTATTACTACTTCGCAATTCT	1057		
QY	236	--LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer	254		
		:::		:::	
Db	1058	CTCGTACTGTCCCATCAACTTCGGACCAAGTTCAGTTCACGGGAC-----	1105		
QY	255	GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu	274		
Db	1106	-----GCAGCAATAAGAGTTCATCATATGAA	1132		
QY	275	MetProIleGluValValSerGlyLeuPheAsnValLysAlaPheGlyAsnAspAla	294		
		:::		:::	
Db	1133	ATGGAGCTTGAGTGAGCTCAGAGCCGTTCAATGTGAGGGAATGCATTTGGGGAAGAGGTT	1192		
QY	295	ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln	314		
Db	1193	GTTCGTGATTAATTCGCGGGTCAGCCCAATGTCCACCGATGAATATGGCTGCTCTTCCAC	1252		
QY	315	ProLeuGlnAsnGlyAlaIleTyrTyrLeuIle	325		
Db	1253	CCTCTTCAACAGGAGCCTCGTACTATCTGATC	1285		

RESULT 3

ADD30297
ID ADD30297

XX
AC

XX
DT 15-JAN-20

XX
DE Plant vield-related vol:mun

Frame Yield-related polynucleotide clone G1584.

ds; transcription factor; transgenic plant; growth rate; senescence;
seed germination rate; plant vigor; seedling vigor.

XX
OS *Arabidopsis thaliana*.

XX
PN WO2003013227-A2.

XX
PD
20-FEB-2003.

XX
PF
09-AUG-2002;

XX
PR 09-AUG-2001: 2001US-0310847P
XX 09-AUG-2002: 2002NO-03023805.

PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P

11-DEC-2001; 2001US-0338692P.
14-JUN-2002; 2002US-00171468.
XX

(MEND-) MENDEL BIOTECHNOLOGY

PI Ratcliffe O, Riechmann JL,
PI Pilonis M

Pilgrim ML, Jiang C, Reuber
PI
PI
Broun PE;

WPI; 2003-248221/24.

P-PSDB; ADD30298.

PT New plant transcription factor polynucleotides and polypeptides, useful

Mon Sep 13 10:30:57 2004

PT in producing transgenic plants with commercially valuable properties,
 PT such as an alteration in a plant growth characteristic, e.g. growth rate
 FT or apomixis.

XX Disclosure; SEQ ID NO 326; 454pp; English.

XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
 CC sequences and their encoded proteins which are especially transcription
 CC factor related cDNA's and proteins. The isolated or recombinant plant
 CC transcription factor polynucleotides and polypeptides are useful in
 CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the cDNAs of the invention.

XX Sequence 1502 BP; 414 A; 390 C; 273 G; 425 T; 0 U; 0 Other;

Alignment Scores: Length: 1502
 Pred. No.: 1,27e-57 Matches: 164
 Score: 623.00 Conservative: 40
 Percent Similarity: 49.64% Mismatches: 83
 Best Local Similarity: 39.90% Indels: 124
 Query Match: 35.58% Gaps: 18
 DB:

US-09-787-737-2 (1-325) x ADD30297 (1-1502)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20
 DB 160 ATGGCTTCTCGATAGACACTGGCCAGCATGTTCAAGTCCAAACCT-----CATCC 213
 QY 21 HisHisGlnHisGluLeuAspThrProSerTyrMethHisTyrSerAsnCysAsnLeu 40
 DB 214 CATCAATGGCAACATGATCACTCCCT-----CTCTGCTTCTGCTTCTCACC 267
 QY 41 SerSerSerPheSerSer-----AspArgLysProAspProLysProArgTyr 56
 DB 268 TCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 327
 QY 57 AsnProLysProGluGlnLeuArgLeuGluSerLysPheAsnSerGlyThrLeuAsn 76
 DB 328 ATTCACAAAGCCAGAGAGATTCGGATCTTCAAGCAATCTTAACTCCGGATGGTGAAT 387
 QY 77 ProProArgGluGluLeuGlnArgLeuArgLeuGluTyrGlyGlnLeuGly 96
 DB 388 CCTCCAGAGAGAG-----ATCAGGCTTCAAGATACGGCCAGTCGGT 432
 QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
 DB 433 GATGCTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 492
 QY 117 ValHisHis----- 119
 DB 493 CTCCTCCACCAACTCTCAAAACACTCTCTCCCTCAAAACCGCAGCCGCGCGCAA 552
 QY 120 -----LysSerProLysMetSerLysLysAspLysThrValilePro--- 133
 DB 553 CCTTCGGCTTCTCTTCT 612
 QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149
 DB 613 AAAACCAAGAACAGACACTAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 669
 QY 150 LeuTyrProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPhePro 169
 DB 670 ATGTTTCCA-----CCGGAACCGCG-----TTTCTCTTCTCCCG 702
 QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183

DB 703 GTCTCCACTGTCGGAGGTTGAAGGTATACCGTCTCATCCCAATTAGGTTCTCTCC 762
 QY GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200
 DB 763 GGTGATATGATTGAGCAACAAACCGCTCCAAAGTGTACCGGACTCTCTGATGAG 822
 QY 201 ValAsnAsnGly-----ValAsnLeuGluThrAsnGlu----- 211
 DB 823 ATCATGAACGGTAGTGTGAGTTATGGAATCATCATCAACACACTTGAAGTGAAGA 882
 QY 211 ----- 211
 DB 883 GTTGAAGAAATGAGATGAAGATGTTGCAACAGCCACAGACTCAGATTTGTTACGCTACC 942
 QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyAspGlyAsnGlyGlyGlyAsn 231
 DB 943 ACTAATCATCAATAGCTTCTTACAC-----AACACACACACACACCAATAC 990
 QY 232 Cys-----PheProPro----- 235
 DB 991 ATCATGCTTCATATATCTCTCCCACTACTTCTACTGCCACACACTATTACTCTCGATTCT 1050
 QY 236 -----LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
 DB 1051 CTCGCTACTGCTCCATCAACTTCGGACCGCTTCAAGTTCACAGCGAC----- 1098
 QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274
 DB 1099 -----GCACGAATAAGAGTTTTTCATCAATGAA 1125
 QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294
 DB 1126 ATGAGCTTGAAGTGAAGTCAAGCCGCTTCAATGTGAGGATGATTTGGGAGAGGTT 1185
 QY 295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln 314
 DB 1186 GTTCTGATTAATTCGCGGCTCAGCCCATGTCCAGATGAATATGGCGTCTCTTCAC 1245
 QY 315 ProLeuGlnAsnGlyAlaIleTyrTyrLeuLe 325
 DB 1246 CCTTTCACACGAGCGCTCTACTATCTGATC 1278
 RESULT 4
 ABK65270
 ID ABK65270 standard; cDNA; 807 BP.
 XX ABK65270;
 AC ABK65270;
 DT 02-JUL-2002 (first entry)
 XX Arabidopsis cDNA encoding a transcription factor #122.
 DE Arabidopsis
 KW Plant; ss; gene; transcription factor; transgenic; agriculture;
 KW metabolic chemical; environmental stress; drought;
 KW microbial disease resistance; herbicide resistance; seed yield;
 KW fruit yield; growth rate; leaf senescence; flower senescence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX WO200215675-A1.
 XX 28-FEB-2002.
 XX 22-AUG-2001; 2001WO-US026189.
 PF 22-AUG-2000; 2000US-0227439P.
 PR 16-NOV-2000; 2000US-00713994.
 PR 18-APR-2001; 2001US-00837944.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 PA (PILG/) PILGRIM M.
 PA (CREE/) CREELMAN R.
 PA (DUBE/) DUBELL A J.

PA (HEAR/) HEARD J.
 PA (JIAN/) JIANG C.
 PA (KEDD/) KEDDIE J.
 PA (ADAM/) ADAM L.
 PA (RATC/) RATCLIFF O.
 PA (REUB/) REUBER J L.
 PA (RIEC/) RIECHMANN J L.
 PA (YUGG/) YU G.
 PA (PINE/) PINEDA O.
 XX
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddle J;
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
 XX
 DR WPI: 2002-292022/33.
 DR P-PSDB: AAU93084.
 XX
 An isolated or recombinant polynucleotide used to produce a transgenic plant.
 XX
 Claim 4; Page 512-513; 941pp; English.
 XX
 The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 polynucleotides encoding an A. thaliana transcription factor

SQ Sequence 807 BP; 234 A; 197 C; 168 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.16e-33	Length:	807
Score:	394.50	Matches:	105
Percent Similarity:	47.82%	Conservative:	38
Best Local Similarity:	35.12%	Mismatches:	101
Query Match:	22.53%	Indels:	55
DB:	6	Gaps:	9

US-09-787-737-2 (1-325) x ABK65270 (1-807)

QY	41	SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro	60
Db	19	TCACATAGCCCATCTCCACTTCCACCAACAGCGCGGACGTTGGTCACTTAACCG	78
QY	61	GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu	80
Db	79	GAGCAATCTTGATCTCGAATCATCTTCAACAGTGTACTGTAAACCCACCAAAAGAT	138
QY	81	GluLeuGlnArgIleArgIleArgLeuGlnGluTrpGlyGlnIleGlyAspAlaAsnVal	100
Db	139	GAACCGGTGAGATAAGAAAGATCTTGAGAAATTCGGTGTGGGAGACGCAACGTC	198
QY	101	PheTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys	120

Db	199	TTCTACTGGTTTCAAAACCGCGGTCAAGATCTCGC	249
QY	121	SerProLysMetSerLysAspLysThrValIleProSerThrAspAlaAspHisCys	140
Db	250	CTTTTAGCAGCCACCGCGCGCCACCTCCATAGAGCTGAAGACCCACGACACATG	309
QY	141	PheGlyPheValAsnGlnGluThrGlyLeuTrpProValGlnAsnGluLeuValVal	160
Db	310	ACGCCCATGAGCATGCATCAA	354
QY	161	ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla	180
Db	355		357
QY	181	PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr	200
Db	358	TTGGGGTTTGA	375
QY	201	ValAsnAsnGlyValAsnLeuGluThrAsn	219
Db	376	AGCAACTTATCAGCTAATTACTTCTTAATGGATCGTCGTCATCTCAATCCCTTCCTTT	435
QY	220	AsnLeu	232
Db	436	TTCTCGGCTCTCTTCTTCAAGTGGTGTGAGACAAACAATGGTATGGAGAATCTC	495
QY	233	PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly	252
Db	496	TTCAAAATGATGGCCATGCAATCTGATATAATCATCAGCAGCAGCATCATGCTCAAT	555
QY	253		267
Db	556	GCTGCATCAGTTTAAACCCATCTGATCAAACTCCACTCCCATACGAACAGAGGG	615
QY	268		286
Db	616	TTTATGACGGTGTATTAACCGAGTTCCTATGGAAGTAACAAAGGAGCAATGACATG	675
QY	287	LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr	306
Db	676	AAAAAATGTCGGTGATGATTCGGTGTACTTCATCTCTGCTCTCTCTCTCTCTCT	735
QY	307	AspGluPheGlyValThrTrpGlnProLeuGlnAsnGlyAlaIleTrpTyrlleu	325
Db	736	CATGAGTTTGGTTTCTTGATGCAATCTTTACACATGACAACTTATTTCTCTGTA	792

RESULT 5

ADD30133
 ID ADD30133 standard; cDNA; 807 BP.

XX AC ADD30133;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related polynucleotide clone G1586.

XX KW ds; transcription factor; transgenic plant; growth rate; senescence;
 seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX FN WO2003013227-A2.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-US025805.

XX PR 09-AUG-2001; 2001US-0310847P.

XX PR 19-NOV-2001; 2001US-0336049P.

XX PR 11-DEC-2001; 2001US-033892P.

XX PR 14-JUN-2002; 2002US-00171468.

(MEND-) MENDEL BIOTECHNOLOGY INC.

Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE, Pilgrim M, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G, Broun PE;

WPI; 2003-248221/24.

P-PSDB; ADD30134.

New plant transcription factor polynucleotides and polypeptides, useful in producing transgenic plants with commercially valuable properties, such as an alteration in a plant growth characteristic, e.g. growth rate or apomixis.

Disclosure; SEQ ID NO 162; 454pp; English.

The invention relates to a number of isolated Arabidopsis thaliana cDNA sequences and their encoded proteins which are especially transcription factor related cDNA's and proteins. The isolated or recombinant plant transcription factor polynucleotides and polypeptides are useful in producing transgenic plants with commercially valuable properties, i.e. modified or altered desirable traits as compared to a reference plant, such as an alteration in a plant growth characteristic, e.g. growth rate, germination rate of seeds, vigor of plants and seedlings, or leaf and flower senescence. Sequence information related to the polynucleotides and polypeptides can also be used in bioinformatic search methods. The transgenic plant is useful for growing a progeny plant from a parent plant. This sequence represents one of the cDNAs of the invention.

Sequence 807 BP; 234 A; 197 C; 168 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,16e-33	Length:	807
Score:	394.50	Matches:	105
Percent Similarity:	47.83%	Conservative:	38
Best Local Similarity:	35.12%	Mismatches:	101
Query Match:	22.53%	Indels:	55
DB:	9	Gaps:	9

US-09-787-737-2 (1-325) x ADD30133 (1-807)

41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60

19 TCACATAGCCATCTCTCCATTCACCGAACCCAGTCCGGGACCGTGTGACCTAACCG 78

61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80

79 GAGCAATCTTGATACGATCGATCCATCTTCAACAGTGTGATCTGTTAACCCCAAAAGAT 138

81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100

139 GAACGGTGGAGATAAGAAAGATGCTTCAGAAATTCGGTGTGTTGGGAGACGCAACGTC 198

101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120

199 TTCTACTGGTTTCAAAACCGACGGTCAAGATCTGCG-----CGGAGACACCGCAG 249

121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140

250 CTTTATAGACGCCACCCAGCCGCCACCTCCATAGGAGTGAAGACACCCAGCACATG 309

141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160

310 ACGGCCATGAGCATGCATCAA-----TATCCTTGCAGCAACACGAGATT----- 354

161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180

355 -----GAT 357

181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200

358 TTGGGGTTTGA-----AGTTGT 375

201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219

376 AGCAACTTATCAGCTAATTAATCTTAAATGATCGTCGTCATCTCAATCCCTTCCTTT 435

220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232

436 TTCTCTCGGCTCTCTCTTCAAGTGGTGGGTGAGAAACAACATGGTATGGAGATCTC 495

233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252

496 TTCAAATGTATGGCCATCATCTCATATAATCATCAGCAGCAGCATCATAGTCAAAAT 555

253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267

556 GCTGCATCAGTTTTTAAACCCATCTCATCAAACTCCAAATCCCAATACGAAACAAGGG 615

268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286

616 TTTATGACGGTGTATTATAACCGAGTTCCTATGGAAGTAAACAAGAGGACCAATAGACATG 675

287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306

676 AAAACAATGTCGGTGAATTCGGTGTACTTCTCTCTGCTTCTCTCTCTCTCTCTCT 735

307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325

736 GATGAGTTTGGTTCTTGTATGATTCATTTTACAACTGGAGCAAACTTATTTCTCTGGTA 792

RESULT 6

ABK65367

ID ABK65367 standard; cDNA; 994 BP.

XX AC ABK65367;

XX XX 02-JUL-2002 (first entry)

XX DE Arabidopsis cDNA encoding a transcription factor #219.

XX KW Plant; ss; gene; transcription factor; transgenic; agriculture;

XX KW metabolic chemical; environmental stress; drought;

XX KW microbial disease resistance; herbicide resistance; seed yield;

XX KW fruit yield; growth rate; leaf senescence; flower senescence.

XX Arabidopsis thaliana.

XX PN W0200215675-A1.

XX XX 28-FEB-2002.

XX XX 22-AUG-2001; 2001WO-US026189.

XX XX 22-AUG-2000; 2000US-0227439P.

XX XX 16-NOV-2000; 2000US-00713994.

XX XX 18-APR-2001; 2000US-00837944.

XX XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX PA (PILG/) PILGRIM M.

XX PA (CREB/) CREELMAN R.

XX PA (DUB/) DUBELL A J.

XX PA (HEAR/) HEARD J.

XX PA (JIAN/) JIANG C.

XX PA (KEDD/) KEDDIE J.

XX PA (ADAM/) ADAM L.

XX PA (RATC/) RATCLIFF O.

XX PA (REUB/) REUBER J L.

XX PA (RIEC/) RIECHMANN J L.

XX PA (YUGG/) YU G.

XX PA (PINE/) PINEDA O.

XX PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;

XX PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;

XX DR WPI; 2002-292022/33.

DR P-PSDB; AAU93181.

XX An isolated or recombinant polynucleotide used to produce a transgenic plant.

PT Claim 4; Page 889-891; 941pp; English.

PS The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 polynucleotides encoding an A. thaliana transcription factor

SQ Sequence 994 BP; 281 A; 209 C; 205 G; 299 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,95e-27 Length: 994
Score: 342.00 Matches: 87
Percent Similarity: 44.22% Conservative: 47
Best Local Similarity: 28.71% Mismatches: 95
Query Match: 19.53% Indels: 75
DB: 6 Gaps: 4

US-09-787-737-2 (1-325) x ABK65367 (1-994)

QY 23 HisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 42
DB 11 AACAAACACCATAGCCCAACCCCATAGTCGCTCACCCCATCC-TCCGCTCCGGT 69
QY 43 SerPheSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGln 62
DB 70 TCCACCTCAGCAGAACCGGTT-----CGGTCCCGATGGTACCTTAACCGAACA 120
QY 63 IleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIle 82
DB 121 ATACTCATCTAGTGTGATCTCCACAGTGTATGGTTAACTCCCAAGAGAGACG 180
QY 83 GlnArgIleArgIleArgLeuGlnGluTyrGlnIleGlyAspAlaAsnValPheTyr 102
DB 181 GTAAGGATAGAAAGATGCTCGAATAATTTCGCCCGGTGGGAGATGCAATGCTCTAT 240
QY 103 TrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerPro 122
DB 241 TGGTTTCAAAACCGCGGTCAAGTCCCGTCGGAGACGCA----- 282
QY 123 LysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGly 142
DB 282 ----- 282
QY 143 PheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValThrGlu 162
DB 282 ----- 282

QY 163 ProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGly 182
DB 283 -----CAGCTACAGGCTGCAGCTGCA 303
QY 183 PheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrValAsn 202
DB 304 GCAGCGGAC-----GCAACCAACCAACTTCTAGCTCTTCTTCTTCTATGGT 351
QY 203 AsnGlyValAsnLeuLeuThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyr 222
DB 352 GTGGATGCTGATATCAAGCAATATGCGCATGGAGATCTCTTAACAATGCT----- 405
QY 223 GlyGlyAspGlyAsnGlyGlyAsnCysPheProLeuThrValProLeuThrIle 242
DB 406 -----GGCCAAATGAGTTACCATGAGCTACTCATCATCATTCATCAA 447
QY 243 AsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGlyAspAsn 262
DB 448 AATCATAGCTCAAAATGTCATCGATTTGTGCCCATCTGATCAAAACTCCAATTTCAA 507
QY 263 ValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValValSerGly 282
DB 508 TACCAACAAGGGGTATACGGTGTATTAACGGAGTTCCGACAGAGTGCAGAGGA 567
QY 283 LeuPheAsnValIleAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGln 302
DB 568 GGAATAGACATGAAGCAACGTTTGGAGAAGATTGGTTTGGTGATTCCTCAGGTGTT 627
QY 303 ProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyr 322
DB 628 CTTCTCTCTACTGATGAGTTGGTTTGTGTCATAGCTTACACATGTTGAGCTTAT 687
QY 323 TyrLeuIle 325
DB 688 TTCCTGGTT 696
RESULT 7
AAS02222
ID AAS02222 standard; cDNA; 1367 BP.
XX AAS02222;
AC AAS02222;
XX
DT 16-JUL-2001 (first entry)
XX
DE Corn Wuschel (WUS) cDNA from clone p0058_chpab57r.
XX
DE Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;
KW transcriptional regulator; proliferation; corn; soybean; wheat; rice;
KW alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;
KW plant transformation; transgenic plant; p0058_chpab57r.
OS Zea mays.
XX
XX Key Location/Qualifiers
XX CDS 202..963
XX /*tag= a
XX /product= "Corn WUSCHEL protein"
XX
XX WO200123575-A2.
XX
PD 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US026648.
XX
XX 30-SEP-1999; 99US-0157216P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon RE, Lowe K;
XX WPI; 2001-258137/26.
DR P-PSDB; AAU01525.

us-09-787-737-2.rng

Mon Sep 13 10:30:57 2004

```
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 18-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139819P.
PR 21-JUN-1999; 99US-0139822P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144322P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149369P.
PR 17-AUG-1999; 99US-0149175P.

PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159395P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Alignment Scores:
Pred. No.: 5,74e-13
Score: 212.50
Percent Similarity: 39.78%
Best Local Similarity: 28.25%
Query Match: 12.14%
DB: 3
Length: 1326
Matches: 76
Conservative: 31
Mismatch: 87
Indels: 75
Gaps: 12

US-09-787-737-2 (1-325) x AAC38580 (1-1326)

Qy 18 AsnAsnHisHisHisGln-----HisGlnLeaspthr 29
Db 283 AACGGCAATGATCACCATCAGCTATATACCCGATCATCAGGACACGATATTCATGAA 342

QY	30	ProSerTyrMethHisTyrSerAsnCysAsnLeuSerSerPheSerSerAspArgIle	49
DB	343	-----CGGAAAAACAACATT	357
QY	50	ProAspProLysPro---ArgTTPAsnProLysProGluGlnIleArgIleLeuGluSer	68
DB	358	CTCTGGCGGCGACTTTGAGATGGAATCCGACGCCAGAGCAGATCACGACGCTAGAAGAG	417
QY	69	IlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArgIleArgIleArg	88
DB	418	CTTTACAGAGCGGAACACGAGCCCGACGACGGAACAGATCCAAACAGATAGCATCTAAG	477
QY	89	LeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTTPheGlnAsnArgLys	108
DB	478	CTTCGTAAATATCGGAGATCGAAGGGAAGACGTTTCTATGTGTTTCAGAAATCATTAAG	537
QY	109	SerArgAlaIleHisIleLysLeuArgValHisHisLysSerProLysMetSerLysLysAsp	128
DB	538	GCTAGAGAGAGACTAAACCGCGC---CGTCGTGAAGGTGGTCTATTATCAACACCAT	594
QY	129	LysThrValIleProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThr	148
DB	595	AAAGAGCTCAAGATTCTATCATCAGGTGTCATCGAGTT-----CATCAGACA	642
QY	149	GlyLeuTyrProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPhe	168
DB	643	AAAGCTCTGCCCA-----TCTTTT	660
QY	169	ProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValVal	188
DB	661	CCACACACAAACCGACCACAG-----	681
QY	189	ProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGlu	208
DB	682	CCACAGCATGAATTAGATCCTCGAGTTTACATAAAGACACAAATGCTAATAATGAAGAT	741
QY	209	-----ThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp	225
DB	742	CATGGGACGACTGAAGATCTCATCAGGGGCATCAGAGGTGGTAAATACGCCACATGG	801
QY	226	GlyAsnGlyGlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSer	245
DB	802	AGAAAT-----CTTGTTACTTGGTCGATACTCAACAA	834
QY	246	GlnGluLys-----ArgAspValGlyLeuSerGly---GlyGluAspValGlyAsp---	261
DB	835	CCGGAAGAGATTAAATATCGAGAAATGTCAACGGAGAGAAGAAACGAGGACAAC	894
QY	262	-----AsnValTyrProValArg	267
DB	895	CGGACTTTAAATCTCTTCCGGTAGG	921
QY	18-OCT-2000	(first entry)	
QY	Arabidopsis thaliana		
QY	Hybridisation assay; genetic mapping; gene expression control;		
QY	protein identification; signal transduction pathway; metabolic pathway;		
QY	promoter; termination sequence; ss.		
QY	Arabidopsis thaliana.		
QY	EP1033405-A2.		
QY	06-SEP-2000.		

PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144008P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144088P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144088P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159291P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161922P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Alignment Scores:		
PR	05-AUG-1999;	99US-0147260P.	Pred. No.:		
PR	06-AUG-1999;	99US-0147303P.	Score:		
PR	06-AUG-1999;	99US-0147416P.	Percent Similarity:		
PR	09-AUG-1999;	99US-0147493P.	Best Local Similarity:		
PR	09-AUG-1999;	99US-0147935P.	Query Match:		
PR	10-AUG-1999;	99US-0148171P.	DB:		
PR	11-AUG-1999;	99US-0148319P.	US-09-787-737-2 (1-325) x AAC47056 (1-1209)		
PR	12-AUG-1999;	99US-0148341P.	Qy		
PR	13-AUG-1999;	99US-0148565P.	Db		
PR	13-AUG-1999;	99US-0148684P.	Qy		
PR	16-AUG-1999;	99US-0149368P.	Db		
PR	17-AUG-1999;	99US-0149175P.	Qy		
PR	18-AUG-1999;	99US-0149426P.	Db		
PR	20-AUG-1999;	99US-0149722P.	Qy		
PR	20-AUG-1999;	99US-0149723P.	Db		
PR	20-AUG-1999;	99US-0149929P.	Qy		
PR	23-AUG-1999;	99US-0149902P.	Db		
PR	23-AUG-1999;	99US-0149930P.	Qy		
PR	25-AUG-1999;	99US-0150566P.	Db		
PR	26-AUG-1999;	99US-0150884P.	Qy		
PR	27-AUG-1999;	99US-0151065P.	Db		
PR	27-AUG-1999;	99US-0151066P.	Qy		
PR	27-AUG-1999;	99US-0151080P.	Db		
PR	30-AUG-1999;	99US-0151303P.	Qy		
PR	31-AUG-1999;	99US-0151438P.	Db		
PR	01-SEP-1999;	99US-0151930P.	Qy		
PR	07-SEP-1999;	99US-0152363P.	Db		
PR	10-SEP-1999;	99US-0153070P.	Qy		
PR	13-SEP-1999;	99US-0153758P.	Db		
PR	15-SEP-1999;	99US-0154018P.	Qy		
PR	16-SEP-1999;	99US-0154039P.	Db		
PR	20-SEP-1999;	99US-0154779P.	Qy		
PR	22-SEP-1999;	99US-0155139P.	Db		
PR	23-SEP-1999;	99US-0155486P.	Qy		
PR	24-SEP-1999;	99US-0155659P.	Db		
			Length:		
			Matches:		
			Conservative:		
			Mismatches:		
			Indels:		
			Gaps:		
			15 LysProCysAsnAsnHisHisGlnHisGluLeuAspThrProSerTyrMetHis 34		
			129 GAGCGCCACAGCATCATCATCAT---CATCAAGCCGACCAAGAAAGCGGCAACAC 185		
			35 TyrSerAsnCysAsnLeuSerSerPheSerSerAspArgPheProAspProLysPro 54		
			186 AACAAACAAAGTCCGGCTCTGGTGGTTACAGTGTCCAG-----ACCAGCAGC 236		
			55 ArgTyrAsnProLysProGluGlnIleArgIleGluSerIle---PheAsnSerGly 73		
			237 AGGTGGACACCGACGACGAGGAGCAATCAAATCTCAAGAACTTTACTACACAACTGCA 296		
			74 ThrIleAsnProProArgGluGluLeuGlnArgIleArgIleGlnGluTyrGly 93		
			297 ATCCGGTCAACCAACAGCCGATCATCCAGAAAGATCATCTCAAGAGCTGAGACAGTTCGGA 356		
			94 GlnIleGlyAspAlaAsnValPheTyrTriPheGlnAsnArgIysSerArgAlaLysHis 113		
			357 AAGATTGAGGCAAGAACGCTTTTACTGGTTCCAGAACCAATAAGGCTCGTGGCGTCAG 416		
			114 LysLeuArgVal----- 117		
			417 AAGAAGAGATTCAACGGAAACAAACATGACCAACCATCTTCATCAACCAACTCGGTTATG 476		
			118 -----HisHisLysSerProLysMetSer 125		

Db	477	ATGGCGGCTAAAGCATATTATCATCTCTACTTACCATCATCACGGTGTTCCTCATGCAG	533
QY	126	LysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheValAsn	145
Db	537	AGACCTGCTAAATTCGGTCAACGTTAAACTTAACCAAGACCATCATCTCTATCATCATTAAC	596
QY	146	AGnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGluProAlaGly	165
Db	597	AAGCCA-----TATCCAGCTTCATTAACGGGAATTTAAATCATGCAAGCTCAGGT	647
QY	166	PheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAsp	185
Db	648	ACTGAATGGTGGTGTGTTAATGCTTCTTAATGGCTACATGAGTAGCCATGCTCTATGATCT	707
QY	186	PheValValProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly---	204
Db	708	-----ATGGAACAAGACTGTTCTAIGAATTACAAACGTAGGTGGAGGATGG	755
QY	205	ValAsnLeuGluThrAsnGluAsn-----PheAspLysIleProAla	218
Db	756	GCAACATCATGATCATCAATTAATCTTACACTTCTTCGATAGA-----GCA	809
QY	219	IleAsnLeuTyrGlyGlyAspGlyAsnGly-----GlyGlyAsnCysPhe	233
Db	810	AAGCCTCTGTTGGTCTAGAGGTCTACAGAAGAAGAAAGTGTGTGGCGATGCTTAT	869
RESULT	11		
ADD30291			
ID	ADD30291	standard; cDNA; 1227 BP.	
XX	AC	ADD30291;	
XX	DT	15-JAN-2004 (first entry)	
XX	DE	Plant yield-related polynucleotide clone G1540.	
XX	KW	ds; transcription factor; transgenic plant; growth rate; senescence;	
XX	KW	seed germination rate; plant vigor; seedling vigor.	
XX	OS	Arabidopsis thaliana.	
XX	PN	WO2003013227-A2.	
XX	PD	20-FEB-2003.	
XX	PF	09-AUG-2002; 2002WO-US025805.	
XX	PR	09-AUG-2001; 2001US-0310847P.	
XX	PR	19-NOV-2001; 2001US-0336049P.	
XX	PR	11-DEC-2001; 2001US-0338692P.	
XX	PR	14-JUN-2002; 2002US-00171468.	
XX	PA	(MEND-) MENDEL BIOTECHNOLOGY INC.	
XX	PI	Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JB;	
XX	PI	Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;	
XX	PI	Broun PE;	
XX	DR	WPI; 2003-248221/24.	
XX	DR	P-PSDB; ADD30292.	
XX	PT	New plant transcription factor polynucleotides and polypeptides, useful	
XX	PT	in producing transgenic plants with commercially valuable properties,	
XX	PT	such as an alteration in a plant growth characteristic, e.g. growth rate	
XX	PS	or apomixis.	
XX	PS	Disclosure; SEQ ID NO 320; 454pp; English.	
XX	CC	The invention relates to a number of isolated Arabidopsis thaliana cDNA	
XX	CC	sequences and their encoded proteins which are especially transcription	
XX	CC	factor related cDNA's and proteins. The isolated or recombinant plant	
XX	CC	transcription factor polynucleotides and polypeptides are useful in	
XX	CC	producing transgenic plants with commercially valuable properties. i.e.	


```

XX SQ Sequence 1338 BP; 328 A; 369 C; 389 G; 252 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.08e-10 Length: 1338
Score: 187.50 Matches: 45
Percent Similarity: 45.89% Conservative: 22
Best Local Similarity: 30.82% Mismatches: 50
Query Match: 10.71% Indels: 29
DB: 4 Gaps: 5

US-09-787-737-2 (1-325) x AAS02219 (1-1338)
QY 25 HisGluIleAspThrProSerTyrMetHisTyr-SerAsn-----CysAsnLeuSerSer 42
Db 124 CACACATATAGCTGAAGCAATATATACCTTCTTAACCTGGCGGTGTAGTGTAGCTGCG 183
QY 43 -----SerPheSer----- 45
Db 184 ATCGTGCAAACTACAGGGGTGTAGTGTATCGATCGCTGATCATATATACCATGGAG 243
QY 46 -----SerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
Db 244 GCGCTAGCGGGCGGTAGCGCTCAAGTGGCGGGGTGAACCTACGGCGGAGCGAGTG 303
QY 64 ArgIleLeuGluSerIlePheAsnSerglyThrIleAsnProProArgGluGluIleGln 83
Db 304 AAGGTCTGACGGAGCTCTTCGCGCGGGCTCGCGACGCCAGCGAGCGAGATCCAG 363
QY 84 ArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
Db 364 CGCATCTCCACCCACCTTCGCGCGCTTCGCGAGGTGGAGAGCAAGCTTCTTACTGG 423
QY 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
Db 424 TTCAGAACCAACAGGCCCGCGAGCGCCAC-----CACCAAGAGAGCGCGC 471
QY 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
Db 472 CGCGCGCGCTCGTGTCC-----TCCCGCAGCGCGGCGAGCGGAGCG 519
QY 144 ValAsnGlnGluThrGly 149
Db 520 AACACGAGGAGACGGC 537

RESULT 14
AAS02227
ID AAS02227 standard; cDNA; 741 BP.
AC AAS02227;
DT 16-JUL-2001 (first entry)
DE Soybean Wuschel (WUS) cDNA from clone ses4d_pk0033_c8.
XX Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;
KW transcriptional regulator; proliferation; corn; soybean; wheat; rice;
KW alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;
KW plant transformation; transgenic plant; ses4d_pk0033_c8.
XX Glycine max.
OS Glycine max.
XX Key Location/Qualifiers
FH 27..665
CDS /*tag= a
FT /product= "Corn WUSCHEL protein"
FT
XX WO200123575-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US026648.
XX

```

30-SEP-1999; 99US-0157216P.
(DUPO) DU PONT DE NEMOURS & CO E I.
Cahoon RE, Lowe K;
WPI; 2001-258137/26.
P-PSDB; AAU01530.

Novel nucleic acid fragments encoding WUS proteins useful for transiently modulating WUS protein level in plant cells, as probes for genetically and physically mapping WUS genes and as markers.

Claim 4; Page 56; 6lpp; English.

The sequence represents cDNA from a clone, which encodes a Wuschel (WUS) protein. The WUS protein plays a key role in initiation and maintenance of the apical meristem of a plant which contains stem cells, allowing for organ formation. WUS encodes a homeodomain protein which is thought to function as a transcriptional regulator. The DNA or its complement is useful for transiently modulating the level of WUS protein in a plant cell and meristem proliferation can be induced through its addition. This involves transforming a plant with WUS DNA to produce a transgenic plant which is grown into a regenerated plant (corn, soybean, wheat, rice, alfalfa, sunflower, canola or cotton). The protein and DNA are useful in the identification of sequences that affect WUS level or enzyme activity, or that encode homologous proteins from the same or other plant species. The polypeptides can then be used to prepare antibodies. Fragments of the DNA may be used as probes to map the genes that they are part of and as markers for traits linked to those genes, which is useful in plant breeding to develop lines with desired phenotypes. The nucleic acid sequences are useful to create transgenic plants in which the WUS polypeptides are present in higher or lower levels than normal or in cell types or developmental stages in which they are not normally found.

Sequence 741 BP; 238 A; 155 C; 156 G; 192 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.51e-10 Length: 741
Score: 187.00 Matches: 62
Percent Similarity: 38.94% Conservative: 19
Best Local Similarity: 29.81% Mismatches: 55
Query Match: 10.68% Indels: 72
DB: 4 Gaps: 9

US-09-787-737-2 (1-325) x AAS02227 (1-741)

```

QY 43 SerPheSerSerAsp-----ArgIleProAspProLysProArgTrp 56
Db 33 AGTCACAGTAGTGATGCTGAGCGGAGAAATGTAGGACTCATTCAGTTTCACGGTGG 92
QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerglyThrIleAsn 76
Db 93 AGTCTCAAAAGGACCAATAGACATGTTAAGAACCTTTTACACGAGGAAATAGGACT 152
QY 77 ProProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGly 96
Db 153 CCCAGCAGCTGAGCAAAATACACAGATTACCTCTAGGCTCAGGGCTTATGGTCACATCGAG 212
QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
Db 213 GGAAGAAATGCTTCTACTGCTGTTTCAAAATCACAAGCTCGTCAAGACAGAGCTG--- 269
QY 117 ValHisHisLysSerProLysMetSerLysLysAspLysThrValIleProSerThr--- 135
Db 270 -----ATGAAGCAACAAACCATTCGATATTCGAATCGC 302
QY 136 -----AspAlaAspHis-----CysPheGlyPheValAsnGln 146
Db 303 TTTCTTCTGCTCCACCCCATTTGCCAAATGTTGCTCGCTCCATATGTTTGCAG 362
QY 147 GluThrGly-----LeuTyrProValGlnAsnAsnGluLeuVal----- 159

```

CC useful in the identification of sequences that affect WUS level or enzyme
CC activity, or that encode homologous proteins from the same or other plant
CC species. The polypeptides can then be used to prepare antibodies.
CC Fragments of the DNA may be used as probes to map the genes that they are
CC part of and as markers for traits linked to those genes, which is useful
CC in plant breeding to develop lines with desired phenotypes. The nucleic
CC acid sequences are useful to create transgenic plants in which the WUS
CC polypeptides are present in higher or lower levels than normal or in cell
CC types or developmental stages in which they are not normally found
XX
SQ Sequence 844 BP; 292 A; 163 C; 166 G; 223 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3 25e-09 Length: 844
Score: 175.50 Matches: 43
Percent Similarity: 51.64% Conservative: 20
Best Local Similarity: 35.25% Mismatches: 34
Query Match: 10.02% Indels: 25
DB: 4 Gaps: 5

US-09-787-737-2 (1-325) x AAS02226 (1-844)

QY 55 ArgTrrAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThr 74
Db 74 CGATGGAGTCTTACAAAGGAGCAATAGACATGTTGGAGAACTTTTACAAAGCAGGAATA 133
QY 75 IleAsnProArgGluGluIleGlnArgIleArgIleLeuGluGlnGlyGln 94
Db 134 AGGACTCCGAGCACTGAGCAATACACAGATTACTCTAGGCTTAGGCTTATGGTTAC 193
QY 95 IleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLys 114
Db 194 ATCGAGGGGAAAAATGCTCTTCTACTGTTTCAAAATCAAAAGCGCGCCAAAGACAGAAG 253
QY 115 LeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIleProSer 134
Db 254 CTCAAG-----CAGAGCAACAAGCATTGCA----- 280
QY 135 ThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrProValGln 154
Db 281 -----TACTGCAATTGCTTCTTCATGCGCTCC-----CACCCCATTTGC 319
QY 155 AsnAsnGluLeuValThrGluPro-----AlaGlyPheLeuPhePro 169
Db 320 CAAAT-----GTTCTCTGCGCTCCATATTGTTTCAAAAGAGAGTGGATTGAGCTTTTAT 373
QY 170 ValHis 171
Db 374 CCTCAC 379

Search completed: September 10, 2004, 09:26:26
Job time : 485 secs

Db 363 CGGAGTGATTACAGTTTATCTTCAACATCGAAGGTGCTTGCAGTGGAGGTATAGT 422
QY 160 ValThrGluProAlaGlyPhe----- 166
Db 423 TCACTGGCCCTTAGGCATGCNAAGATGTTGATGGCATGCAGAGTAGTGAACACCGG 482
QY 167 -----LeuPheProVal----- 170
Db 483 GATTGTAAACCGTGAAGTCTTAACCTCTTCTCTTCATCCAAACCGGCATTTTGAAGAA 542
QY 171 -----HisAsnAspProSerAlaGlnSerAlaPheGlyPheGlyAspPheVal 187
Db 543 AAAACAATCTCAAGTGCCTTCCCTGCTTCACTTCT----- 581
QY 188 ValProValThrGluGluGly 195
Db 582 GTTGTGCTGTTGATGAAGATGGT 605

RESULT 15
AAS02226
ID AAS02226 standard; cdna; 844 BP.
XX
AC AAS02226;
XX
DT 16-JUL-2001 (first entry)
XX
DE Soybean Wuschel (WUS) cDNA from clone scrlc_pk001_d2.
XX
KW Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;
KW transcriptional regulator; proliferation; corn; soybean; wheat; rice;
KW alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;
KW plant transformation; transgenic plant; scrlc_pk001_d2.
XX
OS Glycine max.
XX

Key Location/Qualifiers
FH 8.661
FT CDS
FT /*tag= a
FT /product= "Corn WUSCHEL protein"
FT /partial
FT /note= "No start codon"

XX WO200123575-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US026648.
XX
XX 30-SEP-1999; 99US-0157216P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Cahoon RE, Lowe K;
XX
XX WPI; 2001-258137/26.
XX P-PSDB; AAU01529.
XX

Novel nucleic acid fragments encoding WUS proteins useful for transiently
modulating WUS protein level in plant cells, as probes for genetically
and physically mapping WUS genes and as markers.

Claim 4; Page 55; 61pp; English.
XX
XX The sequence represents cDNA from a clone, which encodes a Wuschel (WUS)
XX protein. The WUS protein plays a key role in initiation and maintenance
XX of the apical meristem of a plant which contains stem cells, allowing for
XX organ formation. WUS encodes a homeodomain protein which is thought to
XX function as a transcriptional regulator. The DNA or its complement is
XX useful for transiently modulating the level of WUS protein in a plant
XX cell and meristem proliferation can be induced through its addition. This
XX involves transforming a plant with WUS DNA to produce a transformed
XX meristem which is grown into a regenerated plant (corn, soybean, wheat,
XX rice, alfalfa, sunflower, canola or cotton). The protein and DNA are

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 09:01:29 ; Search time 91 Seconds
(without alignments)
1981.967 Million cell updates/sec

Title: US-09-787-737-2
Perfect score: 1751
Sequence: 1 MSSNNKWPMSFKPKCNN.....TDFGVGYQLQNGAIYYLI 325

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09787737 @CGN 1.1.69 @runat 01092004 112456.23620 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	6.7	4791	4	US-09-328-352-1416
2	104	5.9	3602	4	US-09-402-929-1
3	104	5.9	6775	4	US-09-402-929-4
4	103.5	5.9	860	4	US-09-031-962D-3
5	102.5	5.9	1404	4	US-09-031-962D-3
6	101.5	5.8	12286	4	US-09-489-039A-619
7	98.5	5.6	2992	1	US-09-322-478-17
8	96.5	5.5	29604	3	US-08-426-236-3
9	96.5	5.5	29604	4	US-08-781-891-207
10	96.5	5.5	4403765	3	US-09-618-166-207
11	96.5	5.5	441529	3	US-09-103-840A-2
12	96	5.5	9829	4	US-09-103-840A-1
					US-09-322-478-19

C	13	93.5	5.3	2260	4	US-09-889-463A-35	Sequence 35, Appl
	14	93	5.3	864	4	US-09-031-962D-1	Sequence 1, Appli
	15	92.5	5.3	936	4	US-09-252-991A-14226	Sequence 14226, A
	16	91	5.2	1318	1	US-08-485-618-100	Sequence 100, App
	17	91	5.2	1318	1	US-08-605-672-100	Sequence 100, App
	18	91	5.2	1318	2	US-08-482-293A-100	Sequence 100, App
	19	91	5.2	1318	2	US-08-943-363-100	Sequence 100, App
	20	91	5.2	1318	3	US-09-193-043-100	Sequence 100, App
	21	91	5.2	1318	4	US-09-688-307A-100	Sequence 100, App
	22	91	5.2	1318	4	US-09-350-259-100	Sequence 100, App
	23	91	5.2	2329	3	US-08-327-219-80	Sequence 100, App
	24	91	5.2	2328	3	US-08-327-219-1	Sequence 80, Appl
	25	91	5.2	3238	3	US-08-927-219-3	Sequence 1, Appli
	26	90.5	5.2	1065	3	US-08-591-685-6	Sequence 3, Appli
	27	90	5.1	2293	1	US-08-604-913B-12	Sequence 6, Appli
	28	90	5.1	3004	1	US-08-276-213-6	Sequence 12, Appl
	29	89.5	5.1	5362	2	US-08-853-310-3	Sequence 6, Appli
	30	89.5	5.1	10480	4	US-09-732-615-13	Sequence 3, Appli
	31	89	5.1	2679	1	US-08-479-328-1	Sequence 1, Appli
	32	89	5.1	2679	1	US-08-761-119-1	Sequence 1, Appli
	33	89	5.1	2679	2	US-08-668-128B-1	Sequence 1, Appli
	34	89	5.1	2679	2	US-08-905-445-1	Sequence 1, Appli
	35	89	5.1	2679	3	US-08-959-625-1	Sequence 1, Appli
	36	89	5.1	2679	3	US-09-008-466-1	Sequence 1, Appli
	37	89	5.1	2679	3	US-08-580-980A-1	Sequence 1, Appli
	38	89	5.1	2679	3	US-09-053-453-1	Sequence 1, Appli
	39	89	5.1	2679	3	US-08-644-116A-1	Sequence 1, Appli
	40	89	5.1	2679	4	US-09-437-858-1	Sequence 1, Appli
	41	89	5.1	4200	1	US-07-841-654B-1	Sequence 1, Appli
	42	89	5.1	4200	1	US-07-946-234A-1	Sequence 1, Appli
	43	89	5.1	4200	1	US-08-123-161A-1	Sequence 1, Appli
	44	89	5.1	4200	1	US-08-483-278-1	Sequence 1, Appli
	45	89	5.1	4200	5	PCT-US93-01560-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-328-352-1416
; Sequence 1416, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1416
; LENGTH: 4791
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1416

Alignment Scores:
Pred. No.: 0.00655

Score: 117.00 Length: 4791
Percent Similarity: 32.90% Matches: 83
Best Local Similarity: 21.50% Conservative: 44
Query Match: 6.68% Mismatches: 139
DB: 4 Indels: 120
Gaps: 20

US-09-787-737-2 (1-325) x US-09-328-352-1416 (1-4791)

QY	8	TriProSerMetPhelySerLysProCysAsnAsnHisHisGlnHisGlu---	26
Db	3664	TGGGAGTCT---TTTAAATCTGCACAGCAATATATCAAGCACTATTTATGAACCT	3720
QY	27	-----	-----
Db	3721	GATAGCTTGTCCACTATTACAGCAGGTTATAGGATTTTATTCAACTGATGAACCT	3780

Mon Sep 13 10:30:58 2004

us-09-787-737-2.rni

```

QY 30 ProSerTyrMetHisTyrSer-----Asn 37
Db 3781 CCTGATTATCAGGAATATCAACAAACCTTATTCTATTATTATAAGACCCAGTATGAAT 3840
QY 38 CysAsnLeuSerSerSer-----PheSerSerAspArg 48
Db 3841 CGTAATCTTGTAAGAGCGTACTGCTTTAGAGAGATTATCACTTACCACTGTAUCAG 3900
QY 49 IleProAspProLysProArgTyrAsnProLysProGluGlnIle---ArgIleLeuGlu 67
Db 3901 GTCGAACACACAAACCATCAACATATCCGAGGAGAGTGTGTATGGAATTTCTCAA 3960
QY 68 SerIlePheAsnSer-----GlyThrIleAsnProProArgGlu 80
Db 3961 GATACATGGCGCGGTGTCGCAAAATTAAGCACTAAATCAAGATAATCCA----- 4011
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db 4012 ---TTTGAACAAATAATTTACGTTTCAA-----GGCAGTATTACGCGTGAACA 4062
QY 101 PheTyrTyrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
Db 4063 GAACGCACTATAACCGTTATCGTTATTATGAACCTCATAGCGCAAGATATGTAAGTAA 4122
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 4123 GACCCG----- 4128
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160
Db 4129 -----ATTGGTTAGAGGGTGTATGAATACCTCTAGTTATGTA 4167
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180
Db 4168 AGTGATCCAAATCAATGGATT-----GATCCAAAGGATTAATAGT----- 4209
QY 181 PheGlyPheClyAsp---PheValValProValThrGluGluGlyMetAlaPheSer 199
Db 4210 TTTAATTATGGTGAAATGTTGGTATTCAGCTAGTGCAGCTAGCTAGCATATCAA 4269
QY 200 ThrValAsnAsn---GlyValAsnLeuGluThrAsnGluAsnPheAsp---LysIlePro 217
Db 4270 GGTCAACAGAAATATGAGTGTATGCTGAAACAGGAGATATGTAATAAAGTGCCA 4329
QY 218 AlaIleAsnLeuTyr-----GlyGlyAspGlyAsnGlyGlyGlyAsnCys 232
Db 4330 CCATTATTGATATGTTGTCATGTAGCGCGGAGGTTTAGAATAGAGTGGCTTTGTT 4389
QY 233 PheProPro-----LeuThrValPro 239
Db 4390 AAAAAATCCATGACTGCGCAATATPATAGTGTGGAAGGATTCGCTATTAAATCCT 4449
QY 240 LeuThrIleAsnGlnSerGlnGluLysArg-----AspValGlyLeuSerGlyGlyGlu 257
Db 4450 GTAGCAAAATCAGTTGCTCAAAATAACAATTTAGCGCTAAAGATCTGGCAGGGCTAGT 4509
QY 258 AspValGlyAspAsnValTyrProVal-----ArgMetThrValPheIleAsn 273
Db 4510 TGTGTAGTGTATATTCACACATACCAAGTTATATCAACAAACACAAATCACTATGGC 4569
QY 274 GluMetProIleGluValValSerGlyLeuPhe-----AsnValLysAla 288
Db 4570 GAGATTACAAATGAATTTGTAAGTGGAGCTTCAGTTACTGTGTGTGTGGAGCTTATGGA 4629
QY 289 AlaPheGlyAsnAspAlaValIleAsnSerPheGlyGlnProIle-----Leu 305
Db 4630 GCAGTACAAACGTAAGTGTTCATATTAGTAATAAGCACTCTGTGTAAGAGGTACGTGG 4689
QY 306 ThrAspGluPheGlyVal 311
Db 4690 GCCTCTGAGTTAGGAGTA 4707

```

RESULT 2

```

US-09-402-929-1
; Sequence 1, Application US/09402929
; Patent No. 641825
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education
; APPLICANT: Toscani, Antonio
; APPLICANT: Hatton, Kimi
; APPLICANT: Reddy, E. P.
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
; NUMBER OF SEQUENCES: 7
; TITLE OF INVENTION: USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402.929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/06896
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-214 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-402-929-1
Alignment Scores:
Pred. No.: 0.137
Score: 104.00
Percent Similarity: 36.18%
Best Local Similarity: 23.08%
Query Match: 5.94%
DB: 4
Length: 3602
Matches: 81
Conservative: 46
Mismatch: 156
Indels: 68
Gaps: 17
US-09-787-737-2 (1-325) x US-09-402-929-1 (1-3602)
QY 10 SerMetPheLysSerLysProCysAsnAsnHisHisHisGlnHisGlu----- 26
Db 869 TCAAAATTCACACAAACCTTGTGGACTATGACCATTTGCAAAACCCAGAAATCAGTTT 928
QY 27 -----IleAspThrProSerTyrMetHisTyrSer-----AsnCys----- 38
Db 929 TACATTCCTGTCAGATCCCTGGGTATCAGTATGTGTCGCTGATGGCAATTTGTTGAA 988
QY 39 ---AsnLeuSerSerPheSerSerAspArgIleProAspProLysProArgTyrTrpAsn 57
Db 989 CATGTTTCAGATCTGCTTTATTCACCAACCTTTGTTGATGAAGATCCT----- 1039
QY 58 ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro 77
Db 1040 GATAAAGAAAAAATAAAGAGCTCGAGTGTCTTCTTATGTCAGCC----- 1087
QY 78 ProArgGluGlnIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97

```

```

Db 1088 ---GAGATGAGTTAGAGGAGGAGGCTTCCACCT---GGAAGCTTTTCTAGC 1141
Qy 98 AlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgVal 117
Db 1142 TGGTCTGTGATCTTCTCATGATGATGATGCTAACACACGACGATGATCTCGAGAA 1201
Qy 118 HisLysSerProLysMetSerLysLysAspLysThrValIleProSerThrAspAla 137
Db 1202 CACACTACTGAGTTTATAGCATGATGATGATGATGATGATGATGATGATGATGATG 1261
Qy 138 AspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGlu 157
Db 1262 CCACAAAGTTCTAGCTGAGAGCAAAATGCTGTGCTCTCTACAG----- 1312
Qy 158 LeuValValThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAla 177
Db 1313 ---ACCATCCAGAAATTCGAGAACTCTGGAATTAATGAAATCGGATCTGTAGCATGG 1369
Qy 178 Gln-----SerAlaPheGlyPheGlyAspPheValVal-----ProValValThr----- 192
Db 1370 AGTGATGTTACCAAGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1429
Qy 193 -----GluGluGlyMetAlaPheSerThrValAsnAsnGly 204
Db 1430 GTTAAATTAATGAGATTCACATAATGAGAGGACCATGGAATGCCAGTTTAAACGTCAGT 1489
Qy 205 ValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGlyGly 224
Db 1490 CTTGTTACTTGAAGGAAAGAAC----- 1513
Qy 225 AspGlyAsnGlyGlyCysPheProProLeuThr----- 237
Db 1514 AGTCGTAATGTTGGAGACAGTGAAGTATTCCTTTAACTCCCAATGCTGCTCAAGTTT 1573
Qy 238 ---ValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGly 256
Db 1574 AGCATCTCTCCACCACTCTCGAAGAAGAAAGAAATTCGAGTGGGTGAGTCTGAGGC 1633
Qy 257 GluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGluMetPro 276
Db 1634 AGTGAGCTGGGATGGCTCACTTACGAGAGTGGTATGATGAGCATCAAAACACACACCA 1693
Qy 277 IleGluValVal-----SerGlyLeuPheAsnValLysAlaAlaPheGlyAsn 292
Db 1694 GTGAAACACTACCAATTTCTCTCTCAGTTTCTTAACT-----ACATGCTCTGGAAT 1747
Qy 293 AspAlaValLeuIleAsn-----SerPhe-----GlyGlnProIleLeuThrAspGluPhe 309
Db 1748 GAACACTTAATATAGAAACCCCTTCTTACATCAACCCCAATTTGTGGCAGAAAGTT 1807
Qy 310 GlyValThrTyrGlnProLeuGlnAsnGlyAla 320
Db 1808 CTCATTACA---ACTCCTCTTCAGAGGAAGCA 1837

RESULT 3
US-09-402-929-4
; Sequence 4, Application US/09402929
; Patent No. 6410825
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education
; APPLICANT: Toscani, Antonio
; APPLICANT: Hattori, Kumi
; APPLICANT: Reddy, B. P.
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/06896
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-214 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-402-929-4

Alignment Scores:
Pred. No.: 0.363 Length: 6775
Score: 104.00 Matches: 81
Percent Similarity: 36.18% Conservative: 46
Best Local Similarity: 23.08% Mismatches: 156
Query Match: 5.94% Indels: 68
DB: 4 Gaps: 17

US-09-787-737-2 (1-325) x US-09-402-929-4 (1-6775)

Qy 10 SerMetPheLysSerLysProCysAsnAsnHisHisHisGlnHisGlu----- 26
Db 4118 TCMAAACTTCAACACAAACCTTGTCCGACTATGGACCAATTGCAAAACCAAGTCAAGTTT 4177
Qy 27 -----IleAspThrProSerTyrMetHisTyrSer-----AsnCys----- 38
Db 4178 TACATTCCTGTTCAGATCCCTGGGTATCAGTATGTCGCTGATGGCAATTGTTGAA 4237
Qy 39 ---AsnLeuSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsn 57
Db 4238 CATGTTCCAGACATTCGCTTTTATTCAGCAACCTTTGTTGATGAGATCCT----- 4288
Qy 58 ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro 77
Db 4289 GATAAAGAAAAAATAAAGGAGCTCGAGTGTCTTCTTATGTCAGCC----- 4336
Qy 78 ProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97
Db 4337 ---GAGAAATGAGTTAGAGGAGAGGCTTCCCACTCAACT---GGAAGCTTTTCTAGC 4390
Qy 98 AlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgVal 117
Db 4391 TGGTCTGTGATGTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 4450
Qy 118 HisHisLysSerProLysMetSerLysLysAspLysThrValIleProSerThrAspAla 137
Db 4451 CACACTACTGAGTTTATAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 4510
Qy 138 AspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGlu 157
Db 4511 CCCACAAAGTTTCTAGCTGAGAGCAAAATGCTGCTGCTGCTTCTCTACAG----- 4561
Qy 158 LeuValValThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAla 177
Db 4562 ---ACCATCCAGAAATTCGAGAAACCTGGAATTAATGAAATCGGATCCTGTAGCATGG 4618

```

```

RESULT 4
US-09-031-962D-3
; Sequence 3, Application US/09031962D
; Patent No. 6350867
; GENERAL INFORMATION:
; APPLICANT: Thomas C. Hart
; APPLICANT: Jennifer A. Price
; TITLE OF INVENTION: Methods and Compositions for Enhancing
; TITLE OF INVENTION: Osseous Growth, Repair, and Regeneration
; FILE REFERENCE: WPU98-18
; CURRENT APPLICATION NUMBER: US/09/031,962D
; CURRENT FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-031-962D-3

Alignment Scores:
Pred. No.: 0.017 Length: 860
Score: 103.50 Matches: 46
Percent Similarity: 40.41% Conservative: 32
Best Local Similarity: 23.83% Mismatches: 59
Query Match: 5.91% Indels: 56
DB: 4 Gaps: 9

US-09-787-737-2 (1-325) x US-09-031-962D-3 (1-860)

Qy 21 HisHisGlnHisGluLeasp-----ThrProSerTyMethHisTyrSer 36
:::|||||:::
Db 193 TACCACCAACCAATTCATCTCAATGGGCTTGACGAGCAGCGGGCGCTTACTCGCCCAAGTCG 252
:::|||||:::

Qy 37 AsnCysAsnLeuSerSerPhe-----
:::|||||:::
Db 253 GAATATATCTACCGAGGCTCTCTACCGGCAATACGGGCGGTATCGGAGCAGCGGCTGCCA 312
:::|||||:::

```

```

RESULT 5
US-09-489-039A-619/c
; Sequence 619, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 619
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-619

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length: 1404
Matches: 52
Conservative: 36
Mismatches: 69
Indels: 65
Gaps: 12

US-09-787-737-2 (1-325) x US-09-489-039A-619 (1-1404)

QY 107 ArgLysSerArgAlaLysHisLys-----LeuArgValHisHis 11
Db 801 CGGCGGAATCTTGGCGGCACCGATCCGACCGCGCTGATCTCGGCTATCACTC 74
QY 120 LysSerPro-----LysMetSerLysLysAspLysThrValIlePro 13
Db 741 CGGCGACGCAATCTTTACTTTGTCCAGGATCTGGCGCACGAAGAATAA---GTGGTTTCAG 68
QY 134 SerThrAspAlaAspHisCysPheGlyPheValAsn-----GlnGluThrGlyLeuTyr 11

```

```

RESULT 5
US-09-489-039A-619/c
; Sequence 619, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 619
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-619

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length: 1404
Matches: 52
Conservative: 36
Mismatches: 69
Indels: 65
Gaps: 12

US-09-787-737-2 (1-325) x US-09-489-039A-619 (1-1404)

QY 107 ArgLysSerArgAlaLysHisLys-----LeuArgValHisHis 11
Db 801 CGGCGGAATCTTGGCGGCACCGATCCGACCGCGCTGATCTCGGCTATCACTC 74
QY 120 LysSerPro-----LysMetSerLysLysAspLysThrValIlePro 13
Db 741 CGGCGACGCAATCTTTACTTTGTCCAGGATCTGGCGCACGAAGAATAA---GTGGTTTCAG 68
QY 134 SerThrAspAlaAspHisCysPheGlyPheValAsn-----GlnGluThrGlyLeuTyr 11

```



```

Db      684  GCCGGCAGCTGGATCACCACCTCTTTTCTGTCAGCGCATTCGACGAGTGGT---TAT 628
Qy      152  ProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis 171
Db      627  CCTTTTGGCATAATACCGGCAGCTTACAGCGCCGACGCTTT----- 583
Qy      172  AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValValProValVal 191
Db      582  ---GATCGGCTGTGGCGCAGGATCGCTTCGCTCACCATCCCGACCGGTGTGTAAGTT 526
Qy      192  ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
Db      525  CAGCAGCCAGCGTGGGACACAGTCTGCTCATATCGCGCAGCGATCTCCAG----- 475
Qy      212  AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGly---Gly 230
Db      474  -----GGCGATGGGAATGTGCGACAGGC 451
Qy      231  AsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys----- 248
Db      450  GTTGGGAAGCGCCCAAGGCCATTGTCTCTCTGGCGATCAAGCCATACCTTGAGTGAGAT 391
Qy      249  -----ArgAspValGlyLeuSerGly-----GlyGluAspVal 259
Db      390  ACGTTTCATCGTGTATCGCGGCTCAAGCAGCGCGCGGCGAACTGCGAGCAGCAGAAATC 331
Qy      260  Gly-----AspAsnValTyrProValArgMetThrValPheIleAsnGluMetPro--- 276
Db      330  GGCATCGCGCAGCGCGCTTTCGCGTCAAGGTGGCCACCACCTTCAATAGTCAGCGCGGC 271
Qy      277  -----IleGluValValSerGlyLeuPhe 284
Db      270  TTTGGCGATCATTCGCGCGCGCAGTCGCGGATAATCTCGACTTCTCTTCTGCTCTTC 211
Qy      285  AsnVal 286
Db      210  AATGTC 205

RESULT 6
US-09-322-478-17
; Sequence 17, Application US/09322478
; Patent No. 6331662
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Voycas, Daniel F.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; FILE REFERENCE: P-1065 ISURF Plant Retroelement
; CURRENT APPLICATION NUMBER: US/09/322,478
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087125
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 12286
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-17

Alignment Scores:
Pred. No.: 1.78 Length: 12286
Score: 101.50 Matches: 76
Percent Similarity: 33.03% Conservative: 33
Best Local Similarity: 23.03% Mismatches: 96
Query Match: 5.80% Indels: 125
DB: 4 Gaps: 16

US-09-787-737-2 (1-325) x US-09-322-478-17 (1-12286)
Qy      20  AsnHisHisGlnHisGluIleAspThrProSerTyrMetHisTyrSer----- 36

```

```

Db      2680 AACCAACATACCAACCCACCATACCAGCACCTCTAGCCAGGTCCGAATCAGCAAGAAAGC 2739
Qy      37  -----AsnCysAsnLeuSerSerPheSerSerAspArgIleProAsp 51
Db      2740 CCACCAAAATAGAGAACTCTGTCATTTTCATCAGAGAGCAAGATCATCATCAAAAGA 2799
Qy      52  -----Prolys 53
Db      2800 GCACGGATGCAGCCATTCCGAATCTAGAGTTCAATGGGCCCACTGGCCCATGACAAAG 2859
Qy      54  ProArg-----TyrAsnProLysProGluGlnIleArg 64
Db      2860 CCAGACGGCCCATAGAACTTTTCGTTGCTAAATGAGAGAGAACACCC---CAGAGNAGG 2916
Qy      65  Ile-----LeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80
Db      2917 ATAAACAGTACTGACTAGAGGC-----AGAGAA 2946
Qy      81  GluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db      2947 GAG---CGCAGGAGGAGGTAAAGTTGAAGAG-----AAGACTGCCAGAGAAAGGAA 2997
Qy      101  PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgVal-HisHisLys 120
Db      2998 GGACAGAGAGACAGAGAGAGAGAGAGAGTGGCAGAGAACCTAAGCTACCAAGAGCC 3057
Qy      120  sSerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db      3058 AGAGAGCAGGAGAACCCAGAGAGAGAGAGAGAGTGGCAGAGTTCACAGGATCTCCCATATC 3117
Qy      140  sPheGlyPheValAsnGlnGluThrGlyLeu-TyrProValGlnAsnAsnGlu----- 157
Db      3118 CTATGCCACCCACCAAGAGAGAGAGAGAGAGTGGCAGGCGTACTTTGCAGTTCTTGGAAAATATCA 3177
Qy      158  -----LeuValValThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerA 176
Db      3178 AAGGGTTAGAAATCACTATGCCATTCGGG-----GAAG 3210
Qy      176  laAlaGlnSerAlaPheGlyPheGlyAspPheValValProValValThrGluGluGlyM 196
Db      3211 CCTTACAGCAGATGCCCTCTACTCCAAATTTATGAAAGACATCTCCACCAAGAGGGA 3270
Qy      196  etAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysI 216
Db      3271 AGTATATTGAC-----AACGAGAATATTGTGTA- 3299
Qy      216  leProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyAsnCys----- 232
Db      3300 -----GGAGGCAATTGCAAGTCGATAA 3321
Qy      233  -----PhePro-----ProLeuThrValProLeuThrI 242
Db      3322 TACAAAGGATTCACCAAGAGAGTTTAAAGACCCCGGAAGTGTACCATCCCTGCCACCA 3381
Qy      242  leAsnGlnSerGlnGluLysArg-----AspValGlyLeuSerGlyGlyGluAspV 259
Db      3382 TTGGGAAGGAGGCGGTAAACAAGGCCCTCATTTGATCTAGAGGCAAGT----- 3428
Qy      259  alGlyAspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGluV 279
Db      3429 -----ATCAATCTGATGCCCTTGTCAA 3450
Qy      279  alValSerGlyLeuPheAsnValLys 287
Db      3451 TGTGCAAAAGAAITGGGAATTTGAAG 3476

RESULT 7
US-08-426-236-3
; Sequence 3, Application US/08426236
; Patent No. 5629188
; GENERAL INFORMATION:
; APPLICANT: Shiba, Kiyotaka

```


Best Local Similarity: 26.49% Mismatches: 90
 Query Match: 5.51% Indels: 75
 DB: 3 Gaps: 17

US-09-787-737-2 (1-325) x US-08-781-891-207 (1-29604)

```

QY 52 ProLysProArgTrpAsnPro---LysProGluGlnIleArg----- 64
Db 22978 CCTAAACACAGAAAGACCCACAAAGATAGAGAACTTCAGACCAATTCTCTTATGAAT 22919
QY 65 IleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArg 84
Db 22918 ATCATCAAAATCTCAATAAAATTCGCTAAGCAATCCAGAACACATTAAGCA 22859
QY 85 IleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPhe 104
Db 22858 ATC---ATCCATCTGACCAAGTAGGTTTATT-----CCAGGATGCGAGGATGTTT 22808
QY 105 GlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLysMet 124
Db 22807 AATATACGAAATCC-----ATCAATGTAATCCATATATAAACAATC 22763
QY 125 SerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheVal 144
Db 22762 AAAGACAAACACCATGATCATC---TCGTTAGATGCAAGAAACATTCGACAACTC 22706
QY 145 AsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAla 164
Db 22705 CAA-----CACCATTCATG-----ATAAAGTGTGGAAAGATCA 22670
QY 165 GlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAlaPheGlyPheGly 184
Db 22669 GGAATTCAGGCCCATCTAATCAATGATAAAGCAATCTTACAGCAA----- 22622
QY 185 AspPheValProValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly 204
Db 22621 -----CCAGTA----- 22616
QY 205 ValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGlyGly 224
Db 22615 GCCAATCATAGTAATGGA-----GAGAAGCTGGAGCAATCCACTAAATCAGGG 22562
QY 225 AspGlyAsnGlyGlyAsnCysPheProLeuThrValProLeu----- 240
Db 22561 ACTAGACAAGC-----TGC-----CCACTTCTCCTACCTTTTCAACATAGTA 22517
QY 241 -----ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
Db 22516 CTTGAAGTATTAGCCAGAGCAATTCGACAGAAAGGAGATCAAGGGGATACAAATT--- 22460
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274
Db 22459 GGAAGAGAGGAAGTCAAAATATCATCTTTTCAGATGATATAGTATATATAGTGAC 22400
QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294
Db 22399 ---CCTAAATTTACAGAGAACTCCTAATC----- 22370
QY 295 ValLeuIleAsnSerPheGlyGln 302
Db 22369 ---CTGATAACAGCTTCGGTGAA 22349

```

RESULT 9

US-09-618-166-207/c

; Sequence 207, Application US/09618166

; Patent No. 6583112

; GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

; Yu, Chang-En

; Oshima, Junko

; Mulligan, John T.

; Schellenberg, Gerald D.

; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

WERNER'S SYNDROME

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-Jul-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McWaters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 207:

SEQUENCE CHARACTERISTICS:

LENGTH: 29604 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 207:

US-09-618-166-207

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
26.5	36.50	29604	71
Percent Similarity:	38.43%	Conservative:	32
Best Local Similarity:	26.49%	Mismatches:	90
Query Match:	5.51%	Indels:	75
DB:	4	Gaps:	17

US-09-787-737-2 (1-325) x US-09-618-166-207 (1-29604)

```

QY 52 ProLysProArgTrpAsnPro---LysProGluGlnIleArg----- 64
Db 22978 CCTAAACACAGAAAGACCCACAAAGATAGAGAACTTCAGACCAATTCTCTTATGAAT 22919
QY 65 IleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArg 84
Db 22918 ATCATCAAAATCTCAATAAAATTCGCTAAGCAATCCAGAACACATTAAGCA 22859
QY 85 IleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPhe 104
Db 22858 ATC---ATCCATCTGACCAAGTAGGTTTATT-----CCAGGATGCGAGGATGTTT 22808
QY 105 GlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLysMet 124
Db 22807 AATATACGAAATCC-----ATCAATGTAATCCATATATAAACAATC 22763
QY 125 SerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheVal 144
Db 22762 AAAGACAAACACCATGATCATC---TCGTTAGATGCAAGAAACATTCGACAACTC 22706
QY 145 AsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAla 164
Db 22705 CAA-----CACCATTCATG-----ATAAAGTGTGGAAAGATCA 22670
QY 165 GlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAlaPheGlyPheGly 184
Db 22669 GGAATTCAGGCCCATCTAATCAATGATAAAGCAATCTTACAGCAA----- 22622
QY 185 AspPheValProValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly 204

```


1594	Db	CTTCTATCA-----							1556
					CCAAAACCTCCTCCTCTCTGTTCTTTCTCCT				
55	QY	-----	-----	-----	-----	ArgTIpAsnProLysPro	60		
1555	Db	CCTTGCTCCTCCATTTCTTTGGCGCTCTTAAACAATATCTCGTTATGGATCCAAACCA					1496		
61	QY	GlulnIleArgIleuGluSerIlePheAsnSerGlyThrIleAsnPro-----					77		
1495	Db	AAGGAAT-----	TCATGGCTTTGACTCAAAAGTGAACCTTTTGGCTT				1448		
78	QY	-----	-----	-----	-----	ProArgIuGlulIcInArg	84		
1447	Db	CTCTTAGTACTGCTCTTTTCGGCGATGTGGTGGCGTTCGCAACCATCTGAAGATCAACAT					1385		
85	QY	IleArgIleArgLeuGlnGluTyr-----				GlyGlnIleGlyAspAlaAsnValPhe	101		
1387	Db	CTTCGTAAACATACCCCAATCTTTTAGTATCAAAGCAATCTCTTACCTTTCACGGCAACAA					1322		
102	QY	TyrTIpPheGlnAsnArgLysSerArgAlaIylHisLysLeuArgVal--HisHisLys					120		
1327	Db	AATGGCATCTCATCATAGAAAGAAAAATCAGAATCTTAACCAATCTCGATCATCACATAT					1263		
121	QY	SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaaspHisCyS					140		
1267	Db	AGTCTCTCTCTCTCTCCCAATTTTTCATATCCCCCAAGTCGATCCGAGTTCGTGT					120		

DEBIT 14

US-09-031-962D-1
; Sequence 1, Application US/09031962D
: Patent No. 6350867

```
, GENERAL INFORMATION:
, APPLICANT: Thomas C. Hart
, APPLICANT: Jennifer A. Price
, TITLE OF INVENTION: Methods and Compositions for Enhancing
, TITLE OF INVENTION: Oodocus Growth, Repair, and Regeneration
, FILE REFERENCE: WTU98-18
, CURRENT APPLICATION NUMBER: US/09/031,962D
, CURRENT FILING DATE: 1998-03-27
, NUMBER OF SEQ ID NOS: 25
, SOURCEAPP: Fast-Seq for Windows Version 3.0
```

Alignment Scores:		
Pred. No.:	0.285	864
Score:	93.00	Matches: 46
Percent Similarity:	40.62%	Conservative: 32
Best Local Similarity:	23.96%	Mismatches: 61
Query Match:	5.31%	Indels: 54
Dg.	4	Gaps: 9

770 00 787 737-3 (1-325) x URS-09-031-962D-1 (1-864)

[illegible]

Db	3188	CTATGCCAATTGGGG-----				-----GAAGCCTTACAGCAGATGC	3220
Qy	181	heGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrV				-----PheP	234
Db	3221	CCCTCTACTCCAATTATGAAAGACATCTCACCAGAGGGGAGTATATTGAC----				-----PheP	234
Qy	201	alAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnL				-----PheP	234
Db	3277	-----AACGAGAAATATTGGGTA-----				-----PheP	234
Qy	221	euTyRGlyGlyAspGlyAsnGlyGlyClyAsnCys-----				-----PheP	234
Db	3295	-----GGAGGCAATTGCAGTGCAGTAAATACAAAGGAAGCTAC				-----PheP	234
Qy	234	ro-----ProLeuThrValProLeuThrIleAsnGlnSerGluG				-----PheP	234
Db	3332	CCAAAGAAGTTTAAAGACCCCGAAGTGTATACCATCCCGTCACCAT-----				-----PheP	234
Qy	247	luLysArgAspValGlyLeuSerGlyGlyGluAspValGlyAspAsnValtyProVala				-----PheP	234
Db	3379	-----GGGAGGAAGCCGTAACAAAGGCCCTCATTTGATCTAA				-----PheP	234
Qy	267	rgMetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal				-----PheP	234
Db	3416	GAGCAAGT-----ATCAATCTGATGCCCTTGTCATATGTCACAAAGAAATTTGGA				-----PheP	234
Qy	287	ys				-----PheP	234
Db	3470	AG				-----PheP	234

RESUIT 13

US-09-889-463A-35/c
Application US/09889463A

Patent No. 6680195
GENERAL INFORMATION:
APPLICANT: Canoon, Rebecca E.
*APPLICANT: Falco, Saverio C.
APPLICANT: Kinney, Anthony J.
APPLICANT: Miao, Guo-Hua
TITLE OF INVENTION: Plant Polyphenol Oxidase
FILE REFERENCE: BB1330
CURRENT APPLICATION NUMBER: US/09/899,463A
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 60/119,590

```

; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 2260
; TYPE: DNA
; ORGANISM: Glycine max
UTS-09-889-463A-35

```

Alignment Scores:			
Pred. No.:	1.1	Length:	2260
Score:	93.50	Matches:	38
Percent Similarity:	36.67%	Conservative:	28
Best Local Similarity:	21.11%	Mismatches:	61
Query Match:	5.34%	Indels:	53
DB:	4	Gaps:	8

722 00 707 737 3 (1-325) x US-09-889-463A-35 (1-2260)

[illegible]

```
QY 79 ArgGluGluIleGlnArgIleArgIle-----ArgLeuGlnGluTyrGly 93
Db 433 CAGCGCGCTCCAGAGAGCCAGTACCTGCGTGGTCCCGCGCGCGAGCTGCGCGC 492
QY 94 GlnIleGly-----AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAla 111
Db 493 CAGTGGGCTCAGCGAGACACAGGTGAATCTGGTTCCAGAACCCGCGTCCAAAGTTC 552
QY 112 LysHisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrVal 131
Db 553 AAG---AAACTCTACAGAAGCGGAGGTGCGCTGGAGCACACTCCCAAT--- 600
QY 132 IleProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyr 151
Db 601 -----AACAGTGATTCCTAGGCGCTGC----- 621
QY 152 ProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPheProValHis 171
Db 622 -----AACTCACC-ACCATCACCGCGCCCTCTGGGACACCTTTCCCACTCCAC 668
QY 172 AsnAspPro-----SerAlaAlaGlnSerAla 180
Db 669 TCGGCGCCCTGCGCGCAGTCAGTGGCCCGCGCGCT 704

RESULT 15
US-09-252-991A-14226
; Sequence 14226, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14226
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14226

Alignment Scores:
Pred. No.: 0.369 Length: 936
Score: 92.50 Matches: 40
Percent Similarity: 39.88% Conservative: 27
Best Local Similarity: 23.81% Mismatches: 62
Query Match: 5.28% Indels: 39
DB: 4 Gaps: 9

US-09-787-737-2 (1-325) x US-09-252-991A-14226 (1-936)
QY 112 LysHisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrVal 131
Db 247 CGGCATGCTGTGGACAACACACCGCGCTGGCGCCAGCGGACCGGACACGACAGTGT 306
QY 132 Ile-----ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGlu 147
Db 307 GTATCCCACTACGGCTGTGTCGGCGGACACACGCGC-----GGTCCCGCGGAC 357
QY 148 ThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAlaGlyPheLeu 167
Db 358 CTACCGCGACACCGTCAACGTGACCATGAGTTGGTGGATGCCGATGCCAGCGCGCTCTA 417
QY 168 Phe-----ProValHisAsnAspProSerAla 176
Db 418 TTCCTGTGCGCTTCTCCTGTCGCGTGGCGAACCTGTCGCGCATGCCATCCATCGCG 477
QY 177 AlaGlnSerAlaPheGlyPheGlyAspPheValProValProValThrGluGluGlyMet 196
```

```
Db 478 ATCAGCGCGCGC-----TTCAGGTCCAGCGGTGGTCCCAACGCGTGC 522
QY 197 AlaPheSerThrVal-----AsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAsp 214
Db 523 GCCTTCGGCACCGCATCAGCAACACGCTATGACCTCGGCACC---CTGAGCTTCGGC 579
QY 215 LysIle-----ProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGly--- 229
Db 580 ACCCTCGGCAACCTGGCGCAGTCCGGTTAAGTCCCGCAGCAGTCCCGGGCGCGGTCCATC 639
QY 230 ---GlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
Db 640 GTCTGACCTGCACTCGGGAATGACCGTCTCGGTGCGCCCTC----- 681
QY 249 ArgAspValGlyLeuSerGlyGly 256
Db 682 ---GACTACGGGTCAATGGCGGC 702
```

Search completed: September 10, 2004, 12:00:29
Job time : 4581 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 10:47:43 ; Search time 564 Seconds
(without alignments)
2899.370 Million cell updates/sec

Title: US-09-787-737-2
Perfect score: 1751
Sequence: 1 MSSNNKWSMPKSKPCNN.....TDFGVYQPLQNGAIYVLI 325

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09787737/runat_01092004.112458.23693/app.query.fasta_1.519
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -WATRX=blousm62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09787737 @CEN 1 1 520 @runat_01092004.112458.23693
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DSLOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

SUMMARIES

ALIGNMENTS

RESULT 1
US-10-225-066A-329
; Sequence 329, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444

1	623	35.6	1502	13	US-10-225-066A-329	Sequence 329, App
2	623	35.6	1502	16	US-10-374-780A-2863	Sequence 2663, App
3	598	34.2	2164	13	US-10-424-599-129204	Sequence 129204, A
4	427	24.4	1961	13	US-10-425-114-27293	Sequence 27293, A
5	405	23.1	2066	17	US-10-437-963-61875	Sequence 61875, A
6	394.5	22.5	807	13	US-09-934-455-243	Sequence 243, App
7	394.5	22.5	807	16	US-10-225-066A-165	Sequence 165, App
8	394.5	22.5	807	13	US-10-374-780A-2665	Sequence 2665, App
9	366	20.9	1524	17	US-10-437-963-59649	Sequence 59649, A
10	354.5	20.2	1215	13	US-10-425-114-12809	Sequence 12809, A
11	354.5	20.2	1218	13	US-10-424-599-97798	Sequence 97798, A
12	342	19.5	994	10	US-09-934-455-489	Sequence 489, App
13	338	19.3	837	16	US-10-260-238-1567	Sequence 1567, App
14	338	19.3	837	17	US-10-437-963-43264	Sequence 43264, A
15	329.5	18.8	921	17	US-10-437-963-59651	Sequence 59651, A
16	328.5	18.8	1367	17	US-10-437-963-59651	Sequence 59651, A
17	307	17.5	789	16	US-10-260-238-5518	Sequence 5518, App
18	304.5	17.4	855	17	US-10-437-963-85739	Sequence 85739, A
19	291.5	16.6	1279	13	US-10-425-114-441	Sequence 441, App
20	212.5	12.1	816	13	US-10-225-066A-331	Sequence 331, App
21	212.5	12.1	816	16	US-10-374-780A-2667	Sequence 2667, App
22	199.5	11.4	1227	13	US-10-412-699B-919	Sequence 919, App
23	199.5	11.4	1227	13	US-10-225-066A-323	Sequence 323, App
24	199.5	11.4	1227	16	US-10-374-780A-301	Sequence 301, App
25	193.5	11.1	2212	17	US-10-744-572-77	Sequence 77, Appl
26	192	11.0	919	17	US-10-437-963-19162	Sequence 19162, A
27	188	10.7	1154	13	US-10-425-114-15010	Sequence 15010, A
28	188	10.7	1289	13	US-10-424-599-113771	Sequence 113771, A
29	188	10.7	2325	17	US-10-744-572-76	Sequence 76, Appl
30	187.5	10.7	1338	17	US-10-744-572-3	Sequence 3, Appl
31	187	10.7	741	17	US-10-424-599-116468	Sequence 116468, A
32	187	10.7	900	13	US-10-424-599-113772	Sequence 113772, A
33	186.5	10.6	981	13	US-10-424-599-113772	Sequence 28807, A
34	185	10.6	1186	13	US-10-425-114-28807	Sequence 23553, A
35	184.5	10.5	731	13	US-10-437-963-23553	Sequence 10583, A
36	183.5	10.5	731	13	US-10-424-599-10583	Sequence 1680, App
37	183.5	10.5	731	13	US-10-412-699B-1680	Sequence 1669, App
38	183.5	10.5	731	16	US-10-424-599-24906	Sequence 35594, A
39	183	10.5	2014	13	US-10-374-780A-1680	Sequence 24906, A
40	181	10.3	998	13	US-10-425-114-35592	Sequence 21484, A
41	179	10.1	423	17	US-10-424-599-129203	Sequence 107605, A
42	177	10.2	423	17	US-10-767-701-21484	Sequence 129203, A
43	176	10.1	688	13	US-10-424-599-129203	Sequence 17, Appl
44	175.5	10.0	844	17	US-10-744-572-17	Sequence 291, App
45	175.5	10.0	870	17	US-10-437-963-291	

PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-12-05
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 1122
SOFTWARE: PatentIn version 3.1
SEQ ID NO 329
LENGTH: 1502
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-225-066A-329

Alignment Scores:
Pred. No.: 5,27e-69 Length: 1502
Matches: 623,00
Conservative: 40
Percent Similarity: 49.64%
Best Local Similarity: 39.90%
Query Match: 35.58%
Indels: 124
Gaps: 18
DB:

US-09-787-737-2 (1-325) x US-10-225-066A-329 (1-1502)

QY 1 MetSerSerAsnLysAsnThrProSerMetPheLysSerLysProCysAsnAsn 20
DB 160 ATGGCTTCTCGAATAGACACTGGCCAGACATGTCACAGTTCACAACTT-----CATCCC 213
QY 21 HisHisHisGlnHisGluLeuAspThrProSerTyMetHisTySerAsnCysAsnLeu 40
DB 214 CATCAATGCCACATGACATCACTCTCT-----CTCTTCCTTCTCTCTCACCGA 267
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTtp 56
DB 268 TCTTCTCTCTTCTTCTCAGATGTGAGTGGAGAGAGTCCAGAGCCAAACCAAGATUG 327
QY 57 AsnProLysProGlnGlnIleLeuGlnSerIlePheAsnSerGlyThrIleAsn 76
DB 328 AATCCAAACCGAGAGAGATCGGATCTTGAAGCAATCTTTAACTCCGGATGGTGAAT 387
QY 77 ProProArgGluGluIleGlnArgIleArgIleArgLeuGlnGlnIleGly 96
DB 388 CTCCAGAGAGAG-----ATCAGGCTTCAAGATACCGCCAAAGTCGT 432
QY 97 AspAlaAsnValPheTyTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
DB 433 GATGCTAACGCTTCTACTGTTCCAAACCGTAAGTCCCGTAGTAACACAACTCCGC 492
QY 117 ValHisHis-----119
DB 493 CTCTCCACACACCACTCCAAACACTCTCTCTCCCTCAACGCAACCGCAGCCGCGCAA 552
QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133
DB 553 CCTTCGGCTTCT 612
QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149
DB 613 AAAAGCAAGAACAAACAACTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 669
QY 150 LeuTyProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPhePro 169
DB 670 ATGTTTCCA-----CCGGAACCGCG-----TTTCTCTTCCCG 702
QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183
DB 703 GTCTCCACTGTGGGAGGGTTTGAAGTATCATCCGCTCTCATCCCAATAGGGTTTCTCTCC 762
QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200

763 GGTGATATGATTGAGCAACAAAACCGGCTCCACGCTGACCGGACTCCTGCTGAGTGAG 822
201 ValAsnAsnGly---ValAsnLeuGluThrAsnGlu-----211
823 ATCATGAACGGTAGTGTGAGTTATGGAATCATCATCAACACACACTGAGTGAGAAAGAA 882
211 -----211
883 GTTGAAGAAATGAGGATGAAGATGTTGCAACAGCCACAGACTCAGATTGTTTACGCTACC 942
212 AsnPheAspLysIleProAlaIleAsnLeuTyGlyGlyAspGlyAsnGlyGlyGlyAsn 231
943 ACTAATCATCAATAGCTCTTCAAC-----AACACACAAACACACATAAAC 990
232 Cys-----PheProPro-----235
991 ATCATGCTTATATATCTCTCCACTTACTTACTGCCACCACTATTACTACTTCCGATTCT 1050
236 ---LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
1051 CTGCTACTGTCCCATCAACTTCGACAGCTTCAAGTTCACGCGGAC-----1098
255 GlyGlyGluAspValGlyAspAsnValTyProValArgMetThrValPheIleAsnGlu 274
1099 -----GCACGAATAAGAGTTTTCATCAATGAA 1125
275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaPheGlyAsnAspAla 294
1126 ATGAGCTTGAAGTGAAGTCAAGACCGTTCATATGTAGGAGATGCAATTTGGGAGAGAGTT 1185
295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyGln 314
1186 GTTCTGATTAATTCGCGGCTCAGCCCATGTGACCGATGAATATGGCGCTCCTCTTCCAC 1245
315 ProLeuGlnAsnGlyValaIleTyTrLeuIle 325
1246 CCTTCTTCAACACGAGGACCTCTACTATCTGATC 1278

RESULT 2
US-10-374-780A-2663
Sequence 2663, Application US/10374780A
Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
APPLICANT: Brown, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omlara
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MEI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468

QY 1 MetSerSerSerAsnLysAsnTyrProSerMetPheLysSerLysProCysAsnAsn 20

Db 627 CTGGGCTTGGCGCGCGCGGTTGTCGGGGCGCGTGGCGGCGCGTCAACGCTGCAGCCT 686
Qy 117 -----
Db 687 TTGATCCCGAACCAATCCAGCTGCTCCCTCGCGCGCGCGCGCGCGGATCTTCCTCGTCC 117
Qy 118 HisHisLysSerProLysMetSerLysLysAspLysThrVal----- 746
Db 747 TCGTCTCTCCGACCGCTCTCTCGGGTCCAGCAAGTCGGCGCGCGCGCGCGGATGTACTAC 131
Qy 132 -----
Db 807 ACGGAGCGCCATGACTTTCAGGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 806
Qy 146 GlnGluThrGlyLeuTyProVal----- 145
Db 867 CAAGCCATCTCCGCGCGCGCTTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 153
Qy 154 -----
Db 927 GCTCCGACGAGCCGCTTTCAGCGGTGGCCACAAAGGCTACTTCTGTGCGCTGCCGAG 926
Qy 163 ProAlaGlyPheLeu----- 162
Db 987 GTCGCGAGCATCTCTGGGGCGCGATACAGGCACAACTGCGCGCGCGCGCGCGCGCGCG 986
Qy 176 AlaAlaGlnSerAla----- 175
Db 1047 GCGGCGTTGCTCGCGGAGCGTTCTTGGGGCTGTACAGGCGCGCGCGCGCGCGCGCGCG 1046
Qy 181 -----
Db 1107 ACTGTCACGACATGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy 186 -----
Db 1167 GCGCGTGTATCATTACAGCACACAGCACAAATACACCGCGCGCGCGCGCGCGCGCGCG 186
Qy 186 -----
Db 1227 AGGACG 1226
Qy 187 -----
Db 1287 AGCG 1286
Qy 198 PheSerThrValAsnAsnGlyValAsn----- 197
Db 1347 GCGAGCGTGGCGCGCGTGGGCTCACATCTTGTGCTGCAAGTACTATGCTGTGTGCTC 1346
Qy 212 AsnPheAspLysIleProAlaIleAsnLeuTyGly----- 211
Db 1407 AACTATGATCTGCTCAAGGACTGGCAGTCCCGGGTCCGCGCGCGCGCGCGCGCGCGCG 1406
Qy 226 GlyAsnGlyGlyGlyAsnCysPheProLeuThrValProLeuThrIleAsnGlnSer 225
Db 1467 AGCACCGCGCGHCA----- 245
Qy 246 GlnGluLysArgAspValGlyLeuSer----- 245
Db 1518 CAGGAGGCGTGTGTACATCTGTCAATCAGGACAGCGTCAAGCGCGCGCGCGCGCGCG 1517
Qy 261 AspAsnValTyProValArgMetThrValPheIleAsnGluMetProIleGluValVal 260
Db 1578 CACAACGTCGCGCGTCCAGGCTC----- 1577
Qy 281 SerGlyLeuPheAsnValLysAlaPheGlyAsnAspAlaValLeuIleAsnSerPhe 280
Db 1602 ----- 1601
Qy 301 GlyGln----- 300
Db 1649 ----- 1649
Qy 313 ----- 313

Db 1650 GCGACCGCGCGCGGCTCGACCTCTGCTCCCGTCCCGTCCGCTGCGGCTGCACAGTG 1709
Qy 314 GlnProLeuGlnAsnGlyAlaIleTyThrLeuIle 325
Db 1710 GAGCGGTCACGACGCGGTCTTTTACTACGTGCTG 1745
RESULT 5
US-10-437-963-61875/c
; Sequence 61875, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 61875
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63265C.1
US-10-437-963-61875
Alignment Scores:
Pred. No.: 7,73e-41 Length: 2066
Score: 405.00 Matches: 143
Percent Similarity: 33.80% Conservative: 50
Best Local Similarity: 25.04% Mismatches: 107
Query Match: 23.13% Indels: 272
DB: 17 Gaps: 18
US-09-787-737-2 (1-325) x US-10-437-963-61875 (1-2066)
Qy 1 MetSerSerSerAsnLysAsnTyrProSerMetPheLysSerLys----ProCysAsnAsn 19
Db 1949 ATGGCTTACCGACAGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1948
Qy 20 AsnHisHisHisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsn 39
Db 1889 CACGACGACGACGACGCTGACATGAAC-----GGCAACGCGCAGC 1851
Qy 40 LeuSerSerSerPhe----- 44
Db 1850 TGTCT 1850
Qy 45 -----SerSerAspArgIleProAspProLysProArgTyr 56
Db 1790 TCTTGTCTCTCTCAGGCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1791
Qy 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76
Db 1730 AACCCGAGCGCGGAGCAGATAGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1731
Qy 77 ProProArgGluGluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGly 96
Db 1670 CCGCGCGCGCGAGATCCCGCGCTCCGCTGTCAGGTCAGGATCGGCGCGCGCGCGCGCGCG 1671
Qy 97 AspAlaAsnValPheTyThrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
Db 1610 GACGCCACGCTCTCTACTGTTCCAGAACCGCAAGTCCCGCACCAAGCAAGCTGCGC 1551
Qy 117 Val-----HisHisLysSer-----ProLysMetSerLysLys 127

[illegible]

```

QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
Db 199 TTCCTCTGGTTCACAAACCGAGTCAAGATCTGC-----CGAGACACCGGACG 249
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 250 CTTTTCAGCAGCCACCGACCGCCACCTCCATAGGAGTGAAGACCAACAGCACATG 309
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGlnLeuValVal 160
Db 310 ACGGCCATGAGCATGCATCAA-----TATCTTCGACCAACACGAGATT----- 354
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
Db 355 -----GAT 357
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
Db 358 TTGGGGTTTGA-----AGTTGT 375
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219
Db 376 AGCAACTTATCAGCTAATTACTTCCCTTAATGGATCGTGCATCTCAATCCCTTCTTT 435
QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys 232
Db 436 TTCCTCGGCCTCTCTTCAAGTGGTGGTGTGAGAACACAAATGGTATGGAGATCTC 495
QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252
Db 496 TTCAAATATGATGGCCATGAATCTCATCATATCATCATCATCATCATCATCATCAT 555
QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267
Db 556 GCTGCATCAGTTTAAACCCATCTGCATCAAACTCCACTCCCAATACCAAGAGGG 615
QY 268 ---MetThrValPheIleAsnGluMetProLeuGluValValSerGlyLeuPheAsnVal 286
Db 616 TTTATGACGGTGTATTAACCGGAGTTCTCTATGAAGTAACAAAGAGGACATAGACATG 675
QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306
Db 676 AAAACAATGTCGGTGATGATCGGTGTTACTTCACTTCTCTGCTCTCTCTCTCTCC 735
QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
Db 736 GATGAGTTTGGTTCTTGTGATGATCTTTTACACATGACAAACTTATTTCTCTGGTA 792

```

RESULT 7

```

US-10-225-066A-165
; Sequence 165, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBEL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: FILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: FINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: W010038-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18

```

```

; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-165

```

Alignment Scores:

```

Pred. No.: 3,84e-40 Length: 807
Score: 394.50 Matches: 105
Percent Similarity: 47.83% Conservative: 38
Best Local Similarity: 35.12% Mismatches: 101
Query Match: 22.53% Indels: 55
DB: 13 Gaps: 9

```

US-09-787-737-2 (1-325) x US-10-225-066A-165 (1-807)

```

QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60
Db 19 TCACATAGCCCATCTCCACTTCCACCGAACAGTCCGGGCACGTTGGTCACTAAACCG 78
QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProArgGlu 80
Db 79 GAGCAAAATCTTGATCTCGAATCCATCTTCAACAGTGTGTCTGTAAACCCACCAAGAT 138
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db 139 GAAACGGTGAGGATAAGAAAGATGCTTGAGAAATTCGGTGTGGTGAGAGACCAACGTC 198
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
Db 199 TTTCTACTGGTTTCAAAACCGGACGTCAGATCTCGC-----CGAGACACCGGACG 249
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 250 CTTTTCAGCAGCCACCGACCGGACCGCCCTCTCATAGGAGTGAAGACCAACAGCATG 309
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGlnLeuValVal 160
Db 310 ACGGCCATGAGCATGCATCAA-----TATCTTCGACCAACACGAGATT----- 354
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
Db 355 -----GAT 357
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
Db 358 TTGGGGTTTGA-----AGTTGT 375
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219
Db 376 AGCAACTTATCAGCTAATTACTTCCCTTAATGGATCGTGCATCTCAATCCCTTCTTT 435
QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys 232
Db 436 TTCCTCGGCCTCTCTTCAAGTGGTGGTGTGAGAACACAAATGGTATGGAGATCTC 495
QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252
Db 496 TTCAAATGATGGCCATGAATCTCATATCATATCATATCATATCATATCATATCAT 555
QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267
Db 556 GCTGCATCAGTTTAAACCCATCTGCATCAAACTCCACTCCCAATACCAAGAGGG 615

```

QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286
 Db 616 TTTATGACGGTGTATTAAACGGAGTTCCTATGGAAGTAAACAAAGGAGCAATAGACATG 675
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306
 Db 676 AAACCAATGTTCCGTGATGATTCGGTGTACTTCACTTCTCTCTCTCTCTCTCTCTCCACT 735
 QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
 Db 736 GATGAGTTTGGTTTCTTGTATGATCTTTATACACATGACAACTATTATTCCTGGTA 792

RESULT 8

US-10-374-780A-2665
 ; Sequence 2665, Application US/10374780A
 ; Publication No. US20040019927A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, Bradley K
 ; APPLICANT: Riechmann, Jose Luis
 ; APPLICANT: Jiang, Cai-Zhong
 ; APPLICANT: Heard, Jacqueline E
 ; APPLICANT: Haake, Volker
 ; APPLICANT: Creelman, Robert A
 ; APPLICANT: Ratcliffe, Oliver
 ; APPLICANT: Adam, Luc J
 ; APPLICANT: Reuber, T. Lynne
 ; APPLICANT: Reddie, James
 ; APPLICANT: Broun, Pierre E
 ; APPLICANT: Pilgrim, Marsha L
 ; APPLICANT: Dubell III, Arnold T
 ; APPLICANT: Pineda, Omaira
 ; APPLICANT: Yu, Guo-Liang
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

FILE REFERENCE: MBI-0047 CIP
 CURRENT APPLICATION NUMBER: US/10/374,780A
 CURRENT FILING DATE: 2003-02-25
 PRIOR APPLICATION NUMBER: 09/837,944
 PRIOR FILING DATE: 2001-04-18
 PRIOR APPLICATION NUMBER: 60/310,847
 PRIOR FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: 09/934,455
 PRIOR FILING DATE: 2001-08-22
 PRIOR APPLICATION NUMBER: 60/336,049
 PRIOR FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/338,692
 PRIOR FILING DATE: 2001-12-11
 PRIOR APPLICATION NUMBER: 10/171,468
 PRIOR FILING DATE: 2002-06-14
 PRIOR APPLICATION NUMBER: 10/225,066
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,067
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,068
 PRIOR FILING DATE: 2002-08-09
 NUMBER OF SEQ ID NOS: 2906
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 2665
 LENGTH: 807
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 FEATURE:
 OTHER INFORMATION: GI586
 US-10-374-780A-2665

Alignment Scores:
 Pred. No.: 3,846-40
 Score: 394.50
 Conservativity: 47.83%
 Best local Similarity: 35.12%
 Indels: 22.53%
 Query Match: 16
 DB: 9

US-09-787-737-2 (1-325) x US-10-374-780A-2665 (1-807)

QY 41 SerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60
 Db 19 TCATATGACCCATCTCCATCCACCGAACAGTCCGGGCGACGTGGTCCACCTAAACCG 78
 QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProArgGlu 80
 Db 79 GAGCAATCTTGATCTCAATCCATCTTCAACAGTGGTACTGTGTAAACCCACCAAGAT 138
 QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
 Db 139 GAAACGGTGGAGTAAAGAAAGATGCTTGAATAATCGTCTGTGGGAGACGCAACAGTC 198
 QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120
 Db 199 TTCTACTGGTTTCAAAACCGAGCGGTCAAGATCTCGC-----CGAGACACCGGAG 249
 QY 121 SerProLysMetSerLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
 Db 250 CTTTTCAGCAGCCACCGCAGCGCCACCTCCATAGAGCTGAAGACCCACCCACCATG 309
 QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuVal 160
 Db 310 ACGGCCATGAGCATGCATCAA-----TATCCTTCAGCAACAACAGATT----- 354
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
 Db 355 -----GAT 357
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
 Db 358 TTGGGGTGTGGA-----AGTTGT 375
 QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnAspAspLysIleProAlaIle 219
 Db 376 AGCAACTTATCAGCTAACTTACTTCTTAATGATCGTGTCTCATCTCAAAATCCCTTCTTT 435
 QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232
 Db 436 TTCTTCGGCTCTCTTCTTCAAGTGGTGGTGGAGAACAAATGGTATGGAGATCTC 495
 QY 233 PheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252
 Db 496 TTCAAAATGTATGGCCATGAATCTGATCATATATCATCAGCAGCATCATAGTCAAAAT 555
 QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267
 Db 556 GCTGCATCAGTTTTTAAACCCATCTGATCAAACTCCCACTCCCAATACGAAAGAGG 615
 QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286
 Db 616 TTTATGACGGTGTATTAAACGGAGTTCCTATGGAAGTAAACAAAGGAGCAATAGACATG 675
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306
 Db 676 AAACCAATGTTCCGTGATGATTCGGTGTACTTCACTTCTCTCTCTCTCTCTCTCCACT 735
 QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
 Db 736 GATGAGTTTGGTTTCTTGTATGATCTTTATACACATGACAACTATTATTCCTGGTA 792

RESULT 9
 US-10-437-963-59649
 ; Sequence 59649, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.


```

: APPLICANT: Barbazuk, Brad
:
: APPLICANT: Li, Ping
:
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
:
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
:
: FILE REFERENCE: 38-21(53221)B
:
: CURRENT APPLICATION NUMBER: US/10/437,963
:
: CURRENT FILING DATE: 2003-05-14
:
: NUMBER OF SEQ ID NOS: 204966
:
: SEQ ID NO 59649
:
: LENGTH: 1524
:
: TYPE: DNA
:
: ORGANISM: Oryza sativa
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: PAT_MFT4530_61253C.1
:
: US-10-437-963-59649

```

Alignment Scores:			
Pred. No.:	4,76e-36	Length:	1524
Score:	366.00	Matches:	128
Percent Similarity:	35.61%	Conservative:	44
Best Local Similarity:	26.50%	Mismatches:	111
Query Match:	20.90%	Indels:	200
DB:	17	Gaps:	16
US-09-787-737-2 (1-325) x US-10-437-963-59649 (1-1524)			
QY	41	SerSerSerPheSerSer	-----AspArgIleProAspProLysProArgTTPAsn 57
Db	76	TCGTGCCCTTCTCGT	CACTGGCGGAGGAGGGTTCCGGACCCGAGCGCGGTGGAC 135
QY	58	ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro	77
Db	136	CGCGGCGGAGCAGATCCGGATCTCTGGAGGCGATCTTCACTCCGGCATGTGTCAACCG 195	
QY	78	ProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluIleGluIleGluIleGlyAsp	97
Db	196	CGCGCGCAGCATCCCGCGCATTCGCATGCACTGTCAGGTGTCAGGAGTACGGCCAGTCCGGCAG 255	
QY	98	AlaAsnValPheTyrTyrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg	116
Db	256	GCCACAGCTCTTCTACTGGTTCCAGAACCGCAAGTCCCGCTCCAGAACAGCTGGCGTCC 315	
QY	116	-----	----- 116
Db	316	GGCGGGAAGCGCGCGCGGGCTGGCTTCGCGCGCAACCGGGGCTCCGCGCGCGCGCG 375	
QY	117	ValHisHisLys	----- 120
Db	376	CGCGACCGGGAGCGCGTGGCGCCGTGTTTCAGCGCGCGCGCCACCAATCTCCCGCGCGCC 435	
QY	121	-----	-----SerProLys 123
Db	436	CAGCGGTGACGCGCAGCAGCAGCTTGTCTCGCCTGTGGCGCGCGCTACTCTCTCTCTCG 495	
QY	124	MetSerLysLysAspLysThr	----- 130
Db	496	TCCTCTCTCCGACCGTTCTCGGGTCCAGCAAGCCTCGAGGGCTACGTTCGACGCAG 555	
QY	131	-----	-----ValIleProSerThrAspAlaAspHis 139
Db	556	GCGATGTCCGTGACGACGCCATGGACCTGCTCTCGCCGCTCCGCGCGGCTGCCACCG 615	
QY	139	-----	----- 139
Db	616	CAGATGCTTATCAAGGCCACCACTGGATCGCCCGCGCGCGCTGCTCCCAAGTGCAC 675	
QY	140	-----	-----CysPheGly 142
Db	676	GGCATCGTCCACACAGCAGCGCGTTCCTCTGTCAGTGGCGCGCAGAGCCCTGCTGTCTG 735	
QY	143	PheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGlu	162
Db	736	GCGCTCGACCTCGGCGCGCCCATCTTTCGGCGCCAGTATCATGCACCTCGCGGTGCCCGCT 795	

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701212767_FLI
US-10-425-114-12809

Alignment Scores:
Pred. No.: 9,89e-35 Length: 1215
Score: 354.50 Matches: 100
Percent Similarity: 43.98% Conservative: 46
Best Local Similarity: 30.12% Mismatches: 105
Query Match: 20.25% Indels: 81
DB: 13 Gaps: 10

US-09-787-737-2 (1-325) x US-10-425-114-12809 (1-1215)
QY 24 GlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 43
DB 90 CAGCATCAAGTGGAGCAAGAGTCCAAAGGCAA-----122
QY 44 PheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
DB 123 ---GGCACCGAGAGAGTGAAGCGGTTCGGTCAAGGTGGACTCCAAAGCCAGAAACAATC 179
QY 64 ArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluLeuGln 83
DB 180 CTCATCTCAGTCCATCTTCACAGTGGCATGTTAACTCCAAAGCCAGAAACCGTC 239
QY 84 ArgIleArgIleArgLeuGluGlnIleGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
DB 240 AGAATAAGGAAGCTTCGAGAAATTCGGGAAATTCGGCGCGTCCAGCCAGCAACGTC 299
QY 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
DB 300 TTCAGAACCGCGCTCCAGATCCCGCCCGCCAGCGCATGATGCAACAGCAGGCT 359
QY 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
DB 360 ACTGCCACCGCCACGCCCGTC-----ACCACCTTTGATCAC-----398
QY 144 ValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValThrGluPro 163
DB 399 -----CCTCAGCCTCAGACTCTTGTAAAT-----428
QY 164 AlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPhe 183
DB 429 GTTGGTGGTGCATTCCTCA-----CATGATCACCACAC-----TTG 464
QY 184 GlyAspPheValValProValProValThrGluGluGlyMetAlaPheSerThrValAsnAsn 203
DB 465 GGAACCTT-----GTTGTGCAACTGAAAGTAGTGTCTTCAACCATGGGTTT 515
QY 204 GlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGly 223
DB 516 GGTGTCTTCTCAATCTCTTATGGTTTCTTGGTTCCTCTCTTCTTCTTCTTCTTGT 575
QY 224 GlyAspGlyAsnGlyGlyGlyAsn-----231
DB 576 GGTGGTGGTGAATTTGGTGGCCACAAACAGGCATGGATGTTTCTTCATCGGTTTCT 635
QY 232 -----CysPhe 233
DB 636 TCTCATCAATAGGGTTTCCCTGATCACCACCACTTCCAGTCTGCTTCATCAGTTGTATC 695
QY 234 ProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeu 253
DB 696 CCTCCTCTGTATCCAAATTTGACCTACACGGCTGGATAT-----734
QY 254 SerGlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsn 273
DB 735 ---GGGGGCCCTAATATTTCAGGATTCATT-----ACAGTGTTCATCAAT 776
QY 274 GluMetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAsp 293
DB 777 GGAATTCGAACCTGAACCTCCAAAGGGCCAAAGACCTCAAAACCGTGTTCGAGAGAT 836
QY 294 AlaValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyr 313
DB 837 GTAATGTTAGTTCATCTCTGGAGTTCATATCCCAATGCAATTCGGTTCCTTGTATG 896
QY 314 GlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
DB 897 CATCACTTCACATGCGCATGCTACTTCTCGTA 932

RESULT 11
US-10-424-599-97798
; Sequence 97798, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 97798
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59325C.1
US-10-424-599-97798

Alignment Scores:
Pred. No.: 9,93e-35 Length: 1218
Score: 354.50 Matches: 100
Percent Similarity: 43.98% Conservative: 46
Best Local Similarity: 30.12% Mismatches: 105
Query Match: 20.25% Indels: 81
DB: 13 Gaps: 10

US-09-787-737-2 (1-325) x US-10-424-599-97798 (1-1218)
QY 24 GlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 43
DB 90 CAGCATCAAGTGGAGCAAGAGTCCAAAGGCAA-----122
QY 44 PheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
DB 123 ---GGCACCGAGAGAGTGAAGCGGTTCGGTCAAGGTGGACTCCAAAGCCAGAAACAATC 179
QY 64 ArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluLeuGln 83
DB 180 CTCATCTCAGTCCATCTTCACAGTGGCATGTTAACTCCAAAGCCAGAAACCGTC 239
QY 84 ArgIleArgIleArgLeuGluGlnIleGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
DB 240 AGAATAAGGAAGCTTCGAGAAATTCGGGAAATTCGGCGCGTCCAGCCAGCAACGTC 299
QY 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
DB 300 TTCAGAACCGCGCTCCAGATCCCGCCCGCCAGCGCATGATGCAACAGCAGGCT 359
QY 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
DB 360 ACTGCCACCGCCACGCCCGTC-----ACCACCTTTGATCAC-----398
QY 144 ValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValThrGluPro 163
DB 399 -----CCTCAGCCTCAGACTCTTGTAAAT-----428
QY 164 AlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPhe 183
DB 429 GTTGGTGGTGCATTCCTCA-----CATGATCACCACAC-----TTG 464
QY 184 GlyAspPheValValProValProValThrGluGluGlyMetAlaPheSerThrValAsnAsn 203
DB 465 GGAACCTT-----GTTGTGCAACTGAAAGTAGTGTCTTCAACCATGGGTTT 515
QY 204 GlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGly 223
DB 516 GGTGTCTTCTCAATCTCTTATGGTTTCTTGGTTCCTCTCTTCTTCTTCTTCTTGT 575
QY 224 GlyAspGlyAsnGlyGlyGlyAsn-----231
DB 576 GGTGGTGGTGAATTTGGTGGCCACAAACAGGCATGGATGTTTCTTCATCGGTTTCT 635
QY 232 -----CysPhe 233
DB 636 TCTCATCAATAGGGTTTCCCTGATCACCACCACTTCCAGTCTGCTTCATCAGTTGTATC 695
QY 234 ProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeu 253
DB 696 CCTCCTCTGTATCCAAATTTGACCTACACGGCTGGATAT-----734
QY 254 SerGlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsn 273
DB 735 ---GGGGGCCCTAATATTTCAGGATTCATT-----ACAGTGTTCATCAAT 776
QY 274 GluMetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAsp 293
DB 777 GGAATTCGAACCTGAACCTCCAAAGGGCCAAAGACCTCAAAACCGTGTTCGAGAGAT 836

```

Db 429 GTGGTGGTCAATTCCA---CATGATCACCACACC-----TTG 464
QY 184 GlyAspPheValProValValThrGluGluMetAlaPheSerThrValAsnAsn 203
Db 465 GGAACACCTT-----GTTGTTGCAACTGAAAGTAGTGTACTTCAACCATGGGTTT 515
QY 204 GlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuThrGly 223
Db 516 GGTGTTCTTCTCATCTTCTTATGAGTTTCTTGGTTCCTCTTCTTCTTCTTCTTCTTGT 575
QY 224 GlyAspGlyAsnGlyGlyGlyAsn----- 231
Db 576 GGTGGTGGTGAATTGGTGGCCCAACAACAAGGATGATGGTTTCTTCTCATCGGTTTCT 635
QY 232 -----CysPhe 233
Db 636 TCTCATCAAAATGGGTTTCCCTGATCACCACACACTTCACTGCTTTCATCAGCTTGTAC 695
QY 234 ProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeu 253
Db 696 CTTCTCTTGTATCCAAATTTGACCTTACCAGGCTGGATAT----- 734
QY 254 SerGlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsn 273
Db 735 ---GGGGGCCCTAAATATTCAGGATTCATT-----ACAGTGTATTATCAAT 776
QY 274 GluMetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAsp 293
Db 777 GGAATTCGACATGAACCTTCAAAAGGGCCCAATAGACTCAAAACCGTTTGGAGAGAT 836
QY 294 AlaValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyr 313
Db 837 GTAAGTTAGTTTCATCTCTGGAGTTCATTCCTCCCAATCCCAATGAATTTGGGTTCTTGATG 896
QY 314 GlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
Db 897 CATACCTTCAACATGGCGATGACTTCTTCTGGTA 932

RESULT 12

US-09-934-455-489
; Sequence 489, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omalra
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 489
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (1)...(711)
; OTHER INFORMATION: G2547
US-09-934-455-489

Alignment Scores:

Pred. No.: 2,89e-33 Length: 994
Score: 342.00 Matches: 87
Percent Similarity: 44.22% Conservative: 47
Best Local Similarity: 28.71% Mismatches: 95
Query Match: 19.53% Indels: 75
DB: 10 Gaps: 4

US-09-787-737-2 (1-325) x US-09-934-455-489 (1-994)

QY 23 HisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 42
Db 11 AACAAACACCATACATCCCAACCCGATAGTCGCTACCCCCATCC-TCCGCTCCGGT 69
QY 43 SerPheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGln 62
Db 70 TCCACCTCAGCAACACCGGTT-----CGTCCCGATGTGTCCTTAACCGGAACA 120
QY 63 IleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIle 82
Db 121 ATACTCATCTAGTCGATCTTCCACAGTGTATGTTAACCTCCCAAGAGAGAG 180
QY 83 GlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyr 102
Db 181 GTAAGGATAAGAAAGATGCTCGAAGAAATTTGGCGGGTGGGAGATGCAATGCTTCTAT 240
QY 103 TrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLysSerPro 122
Db 241 TGGTTTCAAAACCGCGGTCAGGTCGCGTCCGTCGACACACGCA----- 282
QY 123 LysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGly 142
Db 282 ----- 282
QY 143 PheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGlu 162
Db 282 ----- 282
QY 163 ProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGly 182
Db 283 -----CAGCTACAGGCTGCAAGTGC 303
QY 183 PheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrValAsn 202
Db 304 GCAGCGGAC-----GCAACCCACCAACACTTCTAGCTTCTTCTTCTTATGGT 351
QY 203 AsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyr 222
Db 352 GGTGGATGTGATTAATCAAGCAATAGTGGCATGGAGAATCTCTTAAACAATGTCT- 405
QY 223 GlyGlyAspGlyAsnGlyGlyAsnCysPheProLeuThrValProLeuThrIle 242
Db 406 -----GGCCAAATGAGTACCAGTACTTCCAGGCTACTCATCATCATTTATCAA 447
QY 243 AsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluValValSerGly 262
Db 448 AATCATAGCTCAAAATGTCATCGATTTTGTGCCCATCTGATCAAAACTCCAATTTC 507
QY 263 ValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValValSerGly 282
Db 508 TACCACCAAGGGCTATACCGTGTATTAACCGGATTTCCGACAGAGTCCGACGAGCA 567
QY 283 LeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValIleAsnSerPheGlyGln 302
Db 568 GGAATAGACATCAAAAGCAACGTTTGGAGAGATTTGGTTTGGTGCATCTCCAGGTGT 627
QY 303 ProfileLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyr 322
Db 628 CTTCTTCTACTGATGATTTGGTTTGTGATGATAGCTTACACATGGTGGAGCTTAT 687

```

381
343 GGCTCGCGGCTTCTGCTCGCTCGGATGTTCCGCGACGGC-----
192 ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
382 ----GCCGCGGGTACAGTCTCTCGTCTGTCCTCATGTCGTCGCTCC-----423
212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly 228
424 -----TCGCGCGCGTCTGCTGGGATGATGATGGGGGACGTGGAGTACGGGGCGGC 474
229 GlyGlyAsnCysPheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
475 GCGACGACCTGTTCTC-----GCCATCTCGAGCAGATGGGGTAC 513
249 ArgAspValGly-----LeuSerGlyGlyGluAspValGlyAspAsn 262
514 ATGACGCGCGCGGCTGCTGCTCGTGGCGCGCGCGGTCAGCATCAGCAGCAGCAG 573
263 ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu 278
574 CTCTACTACTCGTGTCAACCTGGCAGATGACGGTGTTCATCAGCGAGTGGCGAGG 633
279 ValValSerGlyLeuPheAsnValLysAlaIlePheGlyAsnAspAlaValLeuLeu 298
634 GTGCCAAGGGACCGATCGATCTGATCAATGTTTGGCAGCAGTGTGATGCTGTGCAT 693
299 SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGln 318
694 TCAACGGGTCTCTCTTCCAGCCAAAGAGTACGGCATCTCTCTCCATCTCTCCAGATG 753
319 GlyAlaIleTyrTyrLeuIle 325
754 GCGAGAGCTACTTCTGTT 774

RESULT 14
US-10-437-963-43264
; Sequence 43264, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43264
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46438C.1
US-10-437-963-43264

Alignment Scores:
Pred. No.: 7,136-33 Length: 837
Score: 338.00 Matches: 89
Percent Similarity: 46.34% Conservative: 44
Best Local Similarity: 31.01% Mismatches: 92
Query Match: 19.30% Indels: 62
DB: 16 Gaps: 9

US-09-787-737-2 (1-325) x US-10-260-238-1567 (1-837)

323 TyrLeuIle 325
688 TTCCTGGTT 696

US-10-260-238-1567
; Sequence 1567, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Krepis, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1567
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1567

Alignment Scores:
Pred. No.: 7,136-33 Length: 837
Score: 338.00 Matches: 89
Percent Similarity: 46.34% Conservative: 44
Best Local Similarity: 31.01% Mismatches: 92
Query Match: 19.30% Indels: 62
DB: 16 Gaps: 9

US-09-787-737-2 (1-325) x US-10-260-238-1567 (1-837)

53 LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer 72
58 AGTTCGGTGGAGCGCGGAGCGGAGCAGATCTCATCTGAGTCCATCTTCAACAGC 117
73 GlyThrIleAsnProArgGluGluIleGlnArgIleArgIleArgLeuGlnIleTyr 92
118 GGATGCTGAACCGCGCGGAGCGGAGCGGATCCCGCATCCGAGCTGTCGAGCGCTTC 177
93 GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys 112
178 GCGCGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 237
113 HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle 132
238 GCGCGCGAGCGC-----249
133 ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro 152
250 ---CAGCTGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 306
153 ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis--- 171
307 -----ACTGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 342
172 AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProValVal 191

```

```
Db 58 AGTGGGGGGGAGCGCCGACCGGAGCAGATCTCTGGAGTCATCTTCAACAGC 117
QY GlyThrIleAsnProProArgGluGluIleGlnArgIleArgLeuGlnGluTyr 92
Db 118 GGCATGTGTGAACCGCCGACGAGGACGACCGTCGCGATCCGCAAGCTGCTCGAGCGTTC 177
QY GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys 112
Db 178 GGGCGCGCTCGCGACGCGCAACGCTTCTACTGGTTCAGAACCGCGCTCGCGTCCCGC 237
QY HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle 132
Db 238 CGCGCGCAGCGC-----249
QY ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro 152
Db 250 ---CAGCTGCGAGCGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
QY ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis--- 171
Db 307 -----ACTGCTTCGTCGGTGGCTCGCGCTCGCGCTCGCGCGCGCGCGCGCG 342
QY AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProValVal 191
Db 343 GGCCTCGCGCTTCGCTCGCGTTCGGGATGTTTCGCGCACGCGC-----381
QY ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
Db 382 ---GCCGCGCGGTACAGCTCTCGTCTCGTCTCTCATGTCGCGCGTCC-----423
QY AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly 228
Db 424 -----TCGCCGCGCGTTCGGTGGGATGATGTCGGGAGCGTGAGCTACGGGGCGCG 474
QY GlyGlyAsnCysPheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
Db 475 GCGCAGCAGCTGTTTC-----GCCATCTCGAGGCGAGATGGGTAC 513
QY ArgAspValGly-----LeuSerGlyGlyGluAspValGlyAspAsn 262
Db 514 ATGACGCGCGCGCGGTCTGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 573
QY ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu 278
Db 574 CTCCTACTCTGCTGTCACCTCGGACGATGACGCTGTTTCATCAACGAGTGGCGACGAG 633
QY ValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsn 298
Db 634 GTGCARAGGGACCGATCGATCTGAGATCAATGTTGGCGAGGACGTGATGCTGGTGCAT 693
QY SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsn 318
Db 694 TCAACGGGTGCTCTCTTCAGCAACGAGTACGAGTACGATCTCTCCATTTCTCCAGATG 753
QY GlyAlaIleTyrLeuIle 325
Db 754 GCGGAGAGCTACTTCTCGTT 774
```

RESULT 15

```
US-10-437-963-59651
; Sequence 59651, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 59651
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61255C.1
US-10-437-963-59651
```

```
Alignment Scores:
Pred. No.: 1,02e-31 Length: 921
Score: 329.50 Matches: 61
Percent Similarity: 91.14% Conservative: 11
Best Local Similarity: 77.22% Mismatches: 4
Query Match: 18.82% Indels: 3
DB: 17 Gaps: 1
```

US-09-787-737-2 (1-325) x US-10-437-963-59651 (1-921)

```
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTrpAsn 57
Db 355 TCCTCGCCCTTCTCGTCACTGGGCGGAGGAGGGTTCGCGACCCGGAAGCGCGCGGAAC 414
QY 58 ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro 77
Db 415 CGCGCGCGGAGCAGATCCGATCTCTGGAGGCATCTTCACTCGGCGCATGTGTCAACCG 474
QY 78 ProArgGluGluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97
Db 475 CGCGCGCAGAGATCCGCGCATCCGATGCGAGTGCAGGTACGCGCAGTCTCGCGGAC 534
QY 98 AlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
Db 535 GCCAACGCTTCTACTGTTTCCAGAACCGCAAGTCCCGCTCCAGAACAGAGTGGCG 591
```

Search completed: September 10, 2004, 12:59:25
Job time : 581 secs

This Page Blank (uspto)